REPORT RESUMES

ED 003 587 56

AUTOMATED INSTRUCTION OF REMEDIAL ENGLISH, VOLUME II--RESPONSE LATENCY IN PROGRAMED LEARNING--LATENCY RELATED TO ERROR RATE.

BY- DOROUGH, C. DWIGHT AND OTHERS
HOUSTON UNIV., TEXAS

REPORT NUMBER NDEA-VIIA-551-VOL-2 PUB DATE JUN 61
GRANT OEG-741098
EDRS PRICE MF-$0.27 HC-$6.16 154P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, PROGRAMED MATERIALS, *REMEDIAL COURSES, *ENGLISH INSTRUCTION, *COLLEGE STUDENTS, *GRAMMAR, TEACHING MACHINES, HOUSTON, TEXAS

FIVE CLASSES OF RESPONSE LATENCIES WERE CONSIDERED IN THE STUDY, WHICH EVOLVED FROM CRP 551 (ED 003 586). THEY WERE (1) READING, (2) ANSWERING, (3) SCORING AND TURNING TO THE NEXT FRAME OF MATERIAL, (4) READING PLUS ANSWERING, AND (5) TOTAL FRAME TIME (READING PLUS ANSWERING PLUS SCORING PLUS TURNING). THE SUBJECTS WERE 12 MALES AND 4 FEMALES OBTAINED FROM 6 REMEDIAL ENGLISH CLASSES. IN ADDITION TO THESE SUBJECTS (WHO WORKED IN MATERIALS FROM A MODIFIED TEACHING MACHINE) THERE WERE 129 SUBJECTS WHOSE DATA WERE MADE AVAILABLE AS A RESULT OF REGULAR CLASSROOM ACTIVITY. DATA OBTAINED SUPPORT THE GENERAL VIEW THAT LONGER LATENCIES TEND TO GO WITH ERRORS. THE MEDIAN LATENCY FOR ALL CATEGORIES STUDIED TENDED TO BE GREATER FOR ERROR THAN CORRECT ANSWERS FOR A GIVEN FRAME OF MATERIAL (13 LESSONS OF 1-HOUR LENGTH WERE CONSIDERED). MEDIAN READ-PLUS-ANSWER-LATENCIES OF SMALL GROUPS OF SUBJECTS WERE FOUND TO BE RELATED TO CLASSROOM ERROR RATE OF A LARGE GROUP. (HB)

RESPONSE LATENCY IN PROGRAMMED LEARNING: —
LATENCY RELATED TO ERROR RATE (Ph.d Dissertation).

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

A Dissertation
Presented to

the Faculty of the Department of Psychology
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

B

by
Lloyd O. Brooks
June 1961

ED003587

Acknowledgments

This research was made possible by support from the Office of Education, Department of Health, Education, and Welfare, Title VII of the National Defense Education Act of 1958. Grant No: 741098.01, "Automated Instruction of Remedial English," under the direction of Dr. C. Dwight Dorough, Professor of English and Dr. Martin M. Shapiro, Assistant Professor of Psychology, provided the opportunity for my own research interests to blend with the objectives of a more extensive research effort.

Dr. Martin M. Shapiro, as dissertation committee chairman, has been helpful throughout this study in terms of planning, laboratory space, the loan of crucial equipment, and through help with difficulties as they arose. His help came when invited; I am grateful for the freedom he encouraged in developing this study. Dr. C. Dwight Dorough has kept the needs of this research in view by providing programmed materials when needed and by harmonizing classroom activities with special experimental procedures. Both Miss Jeanette Morgan and Mrs. Patti Welch have been of similar help, and both have contributed thoughtfulness and enthusiasm to an effort to improve education. Discussions with Miss Morgan were important in helping define some issues related to error rate. Through these members of the Remedial English project I became acquainted with several members of the English Department. I look forward to seeing their faith in the automated instruction project justified.

With the exception of Dr. Dorough, the dissertation committee consisted entirely of psychologists: Drs. Hobart G. Osburn, Richard I. Evans, and John F. MacNaughton. While each of these men has contributed to this work, I am especially grateful to Dr. Hobart G. Osburn, a man who can raise standards gently--never jerking them out of reach or overlooking efforts in the right direction.

Dr. Jack Michael was the good teacher in the background, responsible for an initial interest in the possibilities of automated instruction, and a source of important, basic concepts. Without this, the opportunity for the present study would have been less fully exploited.

Dr. Lee Meyerson has provided technical information and suggestions related to the numerous practical problems involved in the preparation of a manuscript. His friendly interest and willingness to consider a wide range of questions has made this study more broadly an educational experience. Mrs. Elizabeth McAuliffe has also helped with practical matters.

My wife, Polly H. Brooks, has helped me throughout this work, contributing to practically every aspect of it.

I am also grateful to Mrs. Evelyn Jones for her help. Mrs. Jones typed the manuscript in final form.

AN "ABSTRACT"

Behavioral measures are needed as guides to revision of programmed materials. Error rate offers less as programs are improved, and revision stops after a phase in which the more generous contributors of error have had disproportionate influence. Error rate should be complemented by the more democratic latency measures. These more sensitive measures may reduce program-revision costs as well as increase the effectiveness of a program.

Study of many current issues may be furthered by measures obtained from each student: assessment of prompting techniques, the role of individual differences, covert- compared with overt-answering, e.g. Latency may define behavior which can be controlled by supplementary reinforcement and, in addition, indicate consequences of attempts at control of program-working behaviors (such as changes in reading rate or in response to error during scoring).

Several latency measures may be obtained by making microswitches part of the mechanism of commercially-available teaching machines. These switches can provide measures during normal student use by completing circuits to various channels of multiple-channel print-out counters, such as the one used in the present study. The complex behavior of working a single frame of a program may be somewhat more amenable to further study when outlined by reading, answering, and scoring latencies. Additional measures could be obtained with further machine modification: e.g., the time taken by the student to

PRECEDING PAGE BLANK-NOT FILMED

score his incorrect answer after the program answer was shown. Non-concealed switches operated by students according to special instructions could provide more detail concerning response to the program.

Sixteen students did some regular remedial English course work on a modified machine. Their data support the general view that longer latencies tend to go with errors. Read, Read-plus-Answer, Score-&-Turn, and Total Frame (but not overt Answer) times all tend to be above S's own average for frames on which S made an error. The median latency for all categories studied tended to be greater for error than correct answer for a given frame of material. (Thirteen lessons of hour length were considered.)

Greatest serendipity entered in the case of end-of-lesson review tests for four lessons. Median Read-plus-Answer latencies of small groups of Ss were found to be related to classroom error rate of a large group (rank correlation coefficients, r_s-values, ranged from 0.58 to 0.83, with N equal to 20 for each). Moreover, when the values for individual Ss were substituted for medians, half of the r_s-values fell in the above range. Amount of material, as a common third factor, was an insufficient explanation for such a strong association between error and latency.

In the case of easier material of a lesson, more of reading time could be accounted for in terms of amount of material.

Table of Contents

Chapter Page

I. The Problem: Measuring and Interpreting Latency 1

II. Greater Guidance from the Final Authority 4

Some Current Issues in Programming 6

Revising a Program 9

Individual Differences as a Possible Complication 13

Restating Some Programming Issues 18

Summary 20

III. The Study of Latencies 22

General Problems in Measurement of Student Performance 23

The Present Study of Latencies 28

Reading 28

Supplementary Reinforcement 29

Answering 33

Reading-plus-Answering 36

Scoring-and-Turning 37

Total Frame time 40

Special problems involved in the use of supplementary reinforcement 40

Summary 42

IV. Method 45

Subjects 45

Apparatus 46

Materials 51

Chapter Page

Procedure 52

V. Results 58

Subject and Lesson Characteristics 58

Data Reduction 59

Latencies for Lessons 70

Intercorrelations of Read, Answer, and S-&-T . 79

First and Second Presentations 83

Latency Related to Error 86

Some Applications to Program Analysis 98

Covert- Compared with Overt-Answering 102

Supplementary Reinforcement 104

Miscellaneous Comparisons 106

VI. Discussion and Summary 108

References 116

Appendix A 123

Appendix B 126

Appendix C 134

List of Tables

Table Page

1 Response Latencies for Subject #6 for Lesson 3-II on First Presentation 60

2 Response Latencies for Subject #6 on First Presentation of Eight Lessons 67

3 Typical Response Latencies for Each Subject on First Presentation 69

4 Typical Response Latencies for Each Subject on Second Presentation 71

5 Response Latencies for Lesson 3 of Unit II on First Presentation 72

6 Typical Response Latencies for Each Lesson for First Presentation 77

7 Typical Response Latencies for Each Lesson for Second Presentation 78

8 Intercorrelations of Read, Answer, and Score-and-Turn Latencies for Subject #6 on First Presentation of Lessons 81

9 Intercorrelations of Read, Answer, and Score-&-Turn Latencies for Lesson and End-of-Lesson Review Test on First Presentation for Two Lessons 82

10 Correlation between First and Second Presentations for Lesson Only 84

11 Summary of Sign Data for Five Categories of Latency on First Presentation 87

Table Page

12 Summary of Sign Tests for Five Categories of Latency on First and Second Presentation 89

13 Summary of Sign Data with Zero-Answer Values Excluded 92

14 Rank Correlations of Classroom Errors with Median Latencies and Amount of Material 93

15 Rank Correlations of R-plus-A Latencies of Individual Subjects with Classroom Errors on Review Test 96

16 Range and Average Response Latencies and Errors Above Average Compared with Errors Below Average on First Presentation 134

17 Range and Average Response Latencies and Errors Above Average Compared with Errors Below Average on Second Presentation 139

18 Review Test Range and Average Response Latencies and E-Scored Errors Above Average Compared with Errors Below Average on First Presentation . . . 141

Chapter I

The Problem: Measuring and Interpreting Latency

An analysis of the time spent on a frame of programmed instructional material may provide a means of determining differences in students' responses to a learning task or differences in programmed materials. The analysis may aid in the determination of parts of a program that are unnecessarily wasteful of student time.

When the time spent on a single frame is broken down, program-working behavior can be described in terms of the time spent on various parts of the process. These times, or latencies, provide a quantitative description of a behavioral process which may lead to its better understanding and control.

The Ss of the present study were remedial English students who did their regular course work at a commercially-built teaching machine. This device was modified only slightly in an effort to time some program-working behaviors: reading, answering, and scoring and turning to the next frame.

Because of the important role played by error rate in current program revising, the relation between various latency measures and error rate received more consideration than any other issue treated. Study of one aspect of this problem may have benefited from classroom-experimental situation similarity. Latency measures obtained from a few Ss in the experimental situation were related to errors made in the regular classroom by a much larger number of students.

The latency measures themselves were of interest as measures. Although latencies measured in future work may involve different operations and refinements, it seemed that a study of reliabilities and intercorrelations, as a function of individual differences and differences in programmed materials, would have useful implications.

A portion of this study relates to the assumption that future research may entail efforts to control certain aspects of the process. Latency data can be considered as helping define a baseline before experimental treatments under controlled conditions are attempted. The use of supplementary reinforcement is a possible means of manipulating behavior for the sake of such study. While latency data provides a means of defining new behavioral topographies, it also points up the extent to which such definitions must be qualified by differences in materials and the individual differences of subjects.

Just how latency-defined topographies should be altered by supplementary reinforcement, in an attempt to improve student learning in an automated teaching situation, is a complex question. It ultimately calls for direct tests in which a number of reinforcement procedures are tried. However, some faulty applications may be prevented by the investigator's becoming familiar, in detail, with what is characteristic of a given student for a given kind of material before applications are attempted. The need for new kinds of operant level-baseline data is acute in the new behavioral territory in which the only contact one has with behavior is the time required for its occurrence.

Chapter II

Greater Guidance from the Final Authority

Teaching machines are being developed along many lines and hold much promise of benefiting education. But, as Skinner (1958) has emphasized, the effectiveness of the machine is largely a matter of the effectiveness of the program of instruction which it presents. And, as Skinner and many others point out, programming, due to its complexity and novelty, tends more toward art than science.

Unfortunately, it is not art by choice but rather art by default. Too much of a program evolves as a result of influences that its creator is not aware of, and thus it is difficult to share the benefits of experience with the next man. Since science is by nature cumulative, there would seem to be much practical gain if art gave ground to the fairly-earned advances of science.

An important starting point has been offered by Skinner (1958). His approach is often contrasted with that of Sidney L. Pressey whom Skinner credits with recognizing, as early as the 1920's, the possibility of machines which not only test, in the case of intelligence and information, but also teach (Skinner, 1958, p. 2). Skinner, who prefers the composed-answer item to the multiple-choice, notes that Pressey had both the problem of cultural inertia and a faulty theoretical emphasis on forgetting as obstacles to overcome (Skinner, 1958, p. 3). Of the two men, Skinner continues to be more fortunate in the yield obtained from his early efforts in this new field. Whatever the reason, conditions simply were not right for a quick and large-scale development from Pressey's early efforts.

Skinner offers constraints for programming based on extrapolations from laboratory research, largely with lower organisms. The efforts of Skinner and his associates have led to classical developments in the field of programmed instruction, such as the Holland-Skinner psychology program discussed by Holland (1960). Skinner's recommendations and suggestions, while sufficient to inspire immediate practical application, still leave a large element of art in the task of programming.

Skinner and his coworkers are now being joined by people with diverse backgrounds and interests in an effort to claim new territory for science at the expense of an unwanted art form. The annotated bibliography of Fry, Bryan, and Rigney (1960) provides a glimpse of the rapid growth in this new field. More recently, Lumsdaine and Glaser (1960) bring many of the important, basic papers together, the contributions of early leaders in programmed learning.

During the transition from art to science, a large element of art may be expected to prevail for many years. A variety of approaches to the writing of programs will be suggested, and new measures will be needed which help evaluate the claims made for each. Many details will need to be known of student response to materials programmed according to both old and new techniques.

Some Current Issues in Programming

Most of the current issues in programming which can be dealt with in terms of error rate can also be dealt with by measures which are at least partial correlates of error rate. Some of the problems involved in prompting are examples. Lumsdaine (1960) points out, as does Israel (Lumsdaine & Glaser, 1960, p. 539), the possibility of simplifying the programming of vanishing prompts, a programming technique suggested by Skinner (1958, p. 11). For Skinner, correct answers are initially made highly probable by a variety of semantic and contextual prompts which are gradually withdrawn or vanished as learning takes place.

Both Lumsdaine and Israel suggest the possible value of making use of some form of physical variation. And Lumsdaine suggests that we consider "the relative merits of having cue clarity governed by the program or the teaching machine (e.g., a fixed intratrial rate of increase) versus having it adjusted by the learner up to the minimum sufficient cue-strength at which he can respond correctly" (Lumsdaine & Glaser, 1960, pp. 538-539).

For definitive tests of the desirability of programmed cue clarity versus learner-adjusted clarity, or of physical dimension cue manipulation versus semantic-contextual prompting, probably a variety of measures will be required. A traditional subject matter achievement test would seem to be limited on the basis of insensitivity, for it tends to reflect the joint influence of several items. Similarly, we shall find that the more straightforward item-by-item error score has its limitations and insensitivity. It can only indicate differences between prompting methods when error is made in one case and not in another. A comparative analysis is furthered neither by the students who make errors in the case of both prompting methods nor by those who make no errors in the case of either.

There are other forms of competition among programming strategies which may require many new measures before we can confidently choose from among rivaling claims. For example, Homme and Glaser (1959)[1] suggest a ruleg programming system, a rule and example system, which has implications for the fading or vanishing of prompts. They claim that they can usually "immediately withdraw the prompt if we have once gotten the behavior out in the presence of the S^D (discriminative stimulus). Similarly, the size-of-step variable fails to concern us in the same way it used to" (Lumsdaine, 1960, p. 495).

---

[1] This reference, and some others which appear later, are cited with the date of first publication so that the reader may be aware of temporal aspects of developments in automated instruction as well as have both the original publication and the secondary source available to him as avenues of approach to the literature. However, in this instance, the 1959 material of Homme and Glaser was read as a paper at the Cincinnati American Psychological Association convention.

When both an original and secondary source are cited in References, only the secondary source was consulted. Exceptions to this are explicitly noted.

They go on to point out that this turns into a problem, for them, of how extensive should the variation in examples be and a collateral problem of systematically breaking down rules into sub-rules.

But what is of concern here is Homme and Glaser's point that changes in programming may change research issues and variables to be investigated (Lumsdaine, 1960, p. 494). In the case of their own programming they noted change regarding both the use of prompts and step-size. Both classes of problems decreased in importance.

New measures are needed which are not only more sensitive than error rate and traditional subject matter achievement measures now employed, but are relatively free of differential commitment to current research issues. These new measures should share with error rate one of its advantages: they should provide some information concerning individual items. For it is here that the revision is made which is to benefit the program.

Another current issue concerns the point at which items should be revised so as to be less difficult. Skinner (1958) has favored a program consisting of rarely missed items on the grounds that correct responses must be emitted in order to be reinforced. While item difficulty may be increased to serve better a motivational purpose, he feels that motivation should be assured in some other way. Skinner grants, as was noted recently by Carr (1959), the possibility that there may be an optimal range of error rate for a given kind of material; and Skinner desires to see this issue treated empirically.

Revising a Program

Largely because of this view favoring a low error rate and the important role played by Skinner in the development of the new field of programmed instruction, error rate has come to be one of the most important determiners of whether or not an item of a program is revised.

This can be related to another view. As we put the student in his proper role, as suggested by Skinner, he becomes the final authority as to the effectiveness or adequacy of a program of instruction. Skinner writes, "An unexpected advantage of machine instruction has proved to be the feedback to the programmer" (Skinner, 1958, p. 18). (The stress on programmer was placed there by Skinner.)

There seems to be a strong tendency, when the student is considered in the role of final authority, to consider, more than any other aspect of this authority's behavior, his errors. The programmer's preoccupation with error can obscure other aspects of the final authority's behavior that also have implications for program revision. However, error rate is a measure serving important, unique functions. It may help point up why new measures are needed if the use made of error data is examined in some detail.

Error rate has much to recommend it, and new measures at the expense of error rate is not at issue. New measures are needed to complement error rate in giving the final authority, the student, greater influence in the way a program develops. As errors pile up they do serve as markers of programming problems which justifiably demand the programmer's attention.

Much progress can be made simply on the basis of error data when the nature of errors is taken into account. Holland (1960), e.g., reports a revision of the Holland-Skinner psychology program that reduced the per cent errors from 20.1 on the first version to 11.0 on the second. With the items rendered less difficult on the second version it was possible for the average student to cover the material of 60 disks, rather than the 48 of the first version, in about one hour less. An interesting side benefit of this revision was that the per cent of items which students had improperly scored decreased from 3.6 to 1.4.

It was suggested earlier that errors direct the programmer to individual items; the content of errors does even more. It can indicate the difficulty due to part of an item. Error content helps the programmer keep what is good about an item as he changes only the defective part. It helps restrict the range of what is guessed about. The programmer can use error content to continue an analysis which goes below the level of an item. While as the programmer goes below item level his practices become more complex and difficult to specify, it is the work at this level which revises the program. Faulty workmanship causes errors to pile up again for an item.

It is this detection of faulty workmanship, more than anything else, which keeps programming from being all art--or science only by way of extrapolations, often tenuous extrapolations. Repair can be repeated, aided, at the level of the item, by error data which is rather objective. A programmer would seem to make slight misjudgement if he were guided only by the punched holes or other marks provided by student self-scoring. The improper scoring by students on the composed-answer Holland-Skinner program is only a minor complication of an objective measure of item difficulty. Meyer (1960) reported a similar finding: only 2.2 per cent improper scoring by eighth-graders on a junior high-school vocabulary program.

It is of considerable practical importance that student self-scoring and expert programmer scoring agree so well. The fact that there is such accuracy in scores made available by students, thus reducing the time required for data analysis, further recommends the use of error rate in program revision. Student self-scoring favors both economy and objectivity, as far as the data the programmer begins with is concerned. Both student self-scoring and the use made of error content by the programmer involve interpretation, however.

Since program revision is first concerned with the detection of items that most need revising, error rate has its greatest value in early revision. At this time there are more errors available to be indicators of differences among items. A few faulty indications, scoring errors, make little difference at such a time.

Nothing is done by the programmer merely on the basis of a number of errors being associated with an item. He simply has his attention restricted to a portion of material which he may be able to change and improve if he can interpret properly error content.

A limitation in the use of error rate in program revision is not that the programmer changes items he should not change as much as it is that items which should be changed may be overlooked. Minor imperfections distributed over many frames of material covering a particular topic may in no instance result in error. Yet these imperfections may blunt or weaken the development of this topic.

When error rate approaches zero, even in the case of an efficient program in which every step adds a sizable increment to the topic under development, further perfecting may result from taking some other aspect of student behavior into account. Scores on an achievement test might rise as a result of changing the sequence of steps in the development of program topics; or items could be more carefully worded so that the same material required less time, for example. After items cease to differ from one another with respect to the number of errors associated with them they still continue to differ with regard to how well they serve their intended function.

Low error rate may be regarded as a necessary but not sufficient goal of effective and efficient programming. Errors afford the programmer less guidance as they become fewer in later stages of program revision. They may come from fewer sources (students).

Individual Differences as a Possible Complication

It is a thesis of the present study that not only are there more aspects of student behavior which the programmer should take into account, but the behavior of each student should influence the final form of the program. Too great an emphasis on error rate alone may obscure what is essentially a need for greater guidance from the final authorities. When errors are the sole indicators of differences between items, the students who make the greater number of errors direct more program revising than do those students who make few errors. It will not be possible to determine how much difference it makes in a program to consider all students, rather than just the error contributors, until measures are developed which reflect the behavior of each student which has implication for program revision. This is a programming issue which requires empirical study.

Skinner & Holland (1958) considered many aspects of student behavior: how they paced themselves (whether or not they put off work at a teaching machine until near a course deadline), under what conditions they tended to cheat, the time required to go through a lesson the first time as well as to repeat portions of the material until all items had been answered correctly, and other aspects such as responses to a questionnaire and voluntary comments.

Skinner and Holland seem to have obtained one measure routinely which was made available by all students and which could have some implication for program revision, at least to the extent that it might indicate which disk needed further revision in order to be more economical of student time. However, they simply seem to use this measure to describe units of their program in terms of time requirements. "The time required to complete a disk depended upon the difficulty. There are also wide individual differences among students. The median time required to go once through the easiest disk (to complete the "first cycle") was 8 minutes with a range of 5 to 13 minutes. The median time for answering all items correctly (by repeating the material) was 10 minutes with a range of 5 to 19 minutes" (Skinner & Holland, 1958, p. 42).

It is a recommendation of the present study that this descriptive use of time be carried much further, that it apply not only to disks or lessons. In fact, such description could profitably occur in the case of very brief behavioral episodes involving part of the sequence of responses to a single frame of material. Here, however, measures based on time are of interest as they bear on the role of individual differences as a complication in automated instruction. Coming from each individual student, they could help indicate--as do no other measures currently in use--just which portions of the material were especially difficult for each student. Such measures directly indicate difficulty in terms of economy of student time. Further study may indicate the conditions under which time also indicates a disposition to err. The lack of basic measures which are sensitive to many aspects of student behavior, to all students, and to small units of programmed material may stand in the way of a clearer understanding of the issues relating to individual differences.

There seem to be on hand few measures which can be immediately applied. Ferster and Sapon (1958) state that aptitude measures, which have been successfully applied in predicting traditional course achievement, worked less well in the case of their programmed material for teaching German (Lumsdaine & Glaser, 1960, p. 182).

Similarly, Meyer (1960) comments on this lack of prediction in the case of her findings concerning her vocabulary program for junior high-school students. "The findings agreed with those of others (Porter, 1959), in that measures of ability (here, the reading test scores would seem to be relevant) showed little relationship to the handling of the materials" (Lumsdaine & Glaser, 1960, p. 236). Meyer goes on to point out that while honors students were superior in terms of both error and posttest scores the "pretest correlated -.52 with the 'gain' score, indicating that, insofar as the test measured what was learned, students who knew little at the start of the experiment gained more from the materials than did the sophisticated students" (Lumsdaine & Glaser, 1960, p. 236).

Keislar (1959) found some individual differences to be important in a program used in teaching arithmetic. Keislar pointed out the high difficulty level of this program. A rank correlation between mental age and posttest gain of 0.52 was obtained (Lumsdaine & Glaser, 1960, p. 431). Keislar considers this possibility: "to accommodate individual differences in ability, two or three versions of the program may be desirable" (Lumsdaine & Glaser, 1960, p. 436).

The Meyer study and the Keislar study both point up the fact that programs can differentially favor certain ability levels; greater prediction seems to be afforded by tests of ability in the case of an unusually difficult program.

It would seem that the greater the participation of all students in guiding the development of an automated instructional procedure--whether it concern program or teaching device--the less it would be unsatisfactory for any subgroup. Continued revision of the Keislar arithmetic program to a certain point on the basis of error data could make it more satisfactory for the students who found it too difficult before revision and it could be done at slight expense to those who found it most adequate before revision. Beyond a certain point, however, such improvement could come to be at the expense of a worsening of the program, for some, due to overprompting. As prompting came to use stronger and stronger cues, the program would come to resemble the traditional textbook more and more.

Skinner has offered the learning principle that the correct response should occur to be reinforced. This would seem to be an optimal point as far as difficulty is concerned, the point of just sufficient stimulus control for responding. Beyond this point the student can afford to become less alert (make fewer observing responses), for his behavior can come under the control of several prompts which have equivalent effects.

While it may be slightly better to subject the better student, in the sense of one who makes fewer errors, to overprompting than the other students to a too difficult program, it may be possible to avoid both extremes.

Items which have been revised toward less difficulty on the basis of error rate may be revised away from overprompting on the basis of latency data. Items which were almost, yet not quite, the occasion for error could result as a compromise in the service of individual differences which was provided by this dual analysis. Although this might not obviate the need for several versions of a program, it could reduce the number needed. If, for practical reasons, there could be only one version of a program, dual analysis could make for a better single version.

The wording of items may make a program in remedial English unsatisfactory for the subgroup of foreign students typically found in such courses. When only error rate is considered the program may seem adequate for such a group. However, when time is considered, another dimension of difficulty may be brought into view. If they simply spend longer on each frame of material, in general, there would be little guidance for accommodating such individual differences. To the extent that the careless wording of a few items cost them several unnecessary minutes such guidance would have practical worth.

Restating Some Programming Issues

Programming should benefit as a result of constraint based on measures from all students, for all frames of material, for important aspects of response to material (reading, observing, answering, etc.), and for dimensions in addition to that of correctness of response (response composition time, e.g.).

Individual differences may become more of a current issue as there is more opportunity for them to appear in program analysis.

New measures, such as those based on time, may challenge or complicate dichotomies such as composed versus multiple choice items and right versus wrong categories of response assessment. The time required to compose an answer may relate to several programming issues.

Student motivation as a function of programming technique is one issue. For Skinner the correct response should occur and be reinforced. Item difficulty which serves to motivate yet violates this principle is not desirable. As items require more time and become more difficult according to this temporal definition, they may increase student motivation (or, at least, one of its correlates: attentiveness). If this difficulty is not so great as to result in incorrect responding, there should be no clash between a learning and a motivating technique. This is just another way of saying that overprompting may result in correct responding at the expense of motivation. A consideration of latency may partially prevent overprompting.

Eventually, empirical study may reveal an ideal range defined in terms of a dual analysis of difficulty. At the lower limit, items may be so easy that reading and answering is essentially reading. At the upper limit, errors may begin to occur. Even if the limits were to be this simple, however, the desired degree of variability of item difficulty from item-to-item would remain as a complication.

Some complication for the student may be expected when, after spending a rather uniform amount of time on each item in a series, he suddenly spends much time on an item he can and does get right.

This condition overlaps with several possible circumstances. The student made some kind of initial mistake which interfered with answering; the program presented awkward or carelessly prepared material; or the aspect of the program requiring much time involved several issues for which a long response time was clearly appropriate.

Occasional frames or items which pose fair questions, which take more time and which most students can answer correctly, may not only render the student more attentive for this kind of item but for those items along the way which require less time. Latency data would be crucial for refined study of this issue.

In revising a program, an error that is associated with a short response time should not have the same status as one associated with a longer time. The item occasioning the former may be revised by doing something to improve student interest and attention several frames ahead of the item that is missed. The latter may suggest work on the particular frame associated with error.

Careless wording of a few frames may have a special cost to subpopulations such as foreign students; imperfections on many frames that are too slight to occasion error may increase non-compensated study effort for all students.

Summary

Error does more than any other aspect of student behavior to keep current programming from being all art, rather than the mixture of art and science it now is.

The programmer is virtually overwhelmed by the amount of material he needs to consider and the care this sometimes entails. Errors provide effective direction to trouble-spots. His efforts to improve a program are more likely to be successful if errors do not pile up again where he has just made his repairs.

Error is necessary but not sufficient; the programmer is likely to stop too soon. As he makes progress in revising a program, errors become fewer and thus, less capable of indicating differences between items. Error rate is an undemocratic measure; all students do not contribute errors equally. As errors become fewer they come from fewer sources. Error rate is a measure which can be and needs to be complemented by other measures. Latency is a measure considered, but other measures should not be overlooked.

Some suggested requirements of desirable complementary measures are that they make future research and programming (program analysis) more effective and more efficient through greater sensitivity. More aspects of final authority, or student, behavior should be considered; all students should help in the perfecting evolution of a program; and, more aspects of a program should come under consideration by way of student reaction.

Chapter III

The Study of Latencies

When it is recognized that the program is essentially a sequence of discriminative stimuli, it becomes easier to recognize that instead of program analysis there is the analysis of program-working behaviors. Programs and teaching machines are of interest because of their association with behavioral processes.

There is nothing particularly new in this view. An early and persisting concern has been whether or not automated instruction is effective. This has meant whether or not there is the desired association between remote behaviors indicative of educational achievement and the use of automated instruction. But, between these remote behaviors (which one might study with a subject-matter achievement test or by way of interview, oral examination, judge or teacher rating, etc.) and error rate, as considered in the previous chapter, there are many behaviors which deserve study as they relate to automated instruction.

From a practical standpoint, remote behavioral consequences may deserve emphasis. However, emphasis on the details of the behavioral process occurring at the time the student receives instruction may do more to favor understanding. A practical side of this latter emphasis may be the better control of remote behaviors indicative of educational achievement through the improved design of automated instructional tasks.

Starting with study of normal use, special modifications could be introduced as normative data, useful to their interpretation, accrued. After it is known how much time is spent on a frame of material, e.g., supplementary reinforcement could be applied to various aspects of program-working behavior in an effort to reduce this time without increasing program error rate, faulty student self-scoring, or errors on an achievement test.

General Problems in Measurement of Student Performance

Because the time spent on a frame of programmed material is so brief (and, perhaps, because the single frame is such a small step in a much larger lesson, unit, and program), many important behaviors are likely to be considered only in terms of some accumulated consequence. Reading rate, e.g., may be of interest in the case of a student who is several lessons behind the class. However, the vantage point for doing something about reading rate would seem to be at the time reading is going on.

The change in focus from educational end results to immediate process pose some interesting new problems in measurement. It is simple enough to obtain total frame times by making a timing system part of the teaching machine mechanism. Similarly, other time measures may be obtained as a result of sensing phase of teaching machine operation. The time from presentation of a frame of new material until S operates a shutter for scoring is another example.

Instructions may be used. S may report by operating a switch that he has completed reading, that he has the solution to a problem in mind, that he is confident his answer is correct, etc.

New features may be added to the teaching machine. The pressure from writing or the operation of a keyboard may indicate the beginning of answering, answer-shutter operation the end of answering, etc. S could wear special headgear that involved a light. As he viewed a frame of material his light could act on a photoelectric cell. This time should contain a large element of reading time. When S turned toward the answer-space another photocell circuit could be activated, timing predominately an active, overt answering phase. Special shutters could be added to control visual access to material or to control access to the answer-space.

Regardless of technique there is the problem of deciding on what behavior to measure. Some behaviors are more likely to form a unit than others. If one measures from the beginning to the end of a behavioral unit (or operant) which is maintained under (1) the precise control of a discriminative stimulus, (2) a stable reinforcing contingency, and (3) a consistently effective reinforcing event, there should be two advantages. One, relative to other points of measurement the measure taken from beginning to end should be more stable, or reliable (lawful or regular). Two, it should be easier to relate to other behavioral measures, easier to label, identify, or classify.

Accordingly, relative to other behavioral segments capable of measurement, the sequence beginning with the student reading a new frame of material and ending with his checking his answer against that of the program may be a good starting point. Here there is rather well-defined stimulus control (new frame as S^D), probably adequate knowledge-of-results reinforcement (especially in the case of a program with low error rate so that there is much student and program-answer agreement), and a simple reinforcement contingency (the student writes his answer and without further ado compares it with that of the program).

It may be a better starting point than behavioral sequences which are shorter, or those which are longer, for it forms a unit which can be maintained as a unit. It begins with a clear-cut stimulus event which serves as the S^D and ends with a consistent class of event which is likely to be reinforcing.

Whether measurement is accomplished through normal machine operation, instruction, or features added for the sake of special measures, the behaviors which are being measured (and those which are part of the measurement operation as a result of special instructions) need to be considered in terms of stimulus control, contingency, and reinforcer.

The scoring of errors by the student may serve as an example; a certain amount of detail seems necessary. Suppose a student were asked to operate a scoring mechanism which was located out of sight and which required considerable postural adjustment. Accuracy of scoring could be reduced through faulty stimulus control. The presence of an error and the obvious presence of a scoring mechanism could not combine to set the occasion for error scoring. There would be only error to serve this function.

However, the aversiveness of effort resulting from a cumbersome scoring procedure could account, even more, for faulty scoring. In the presence of an error, S's scoring behavior would always entail some unpleasantness. Not scoring might or might not have a less desirable consequence. In any event, the consequence of not scoring would be delayed longer than the aversiveness of radical postural adjustments and scoring effort. Under these conditions, the aversiveness of scoring behavior could result in faulty reading of answers or forgetfulness.

This would be especially true if S were told only that his errors might eventually be considered for the sake of revising the program. Faulty scoring could be expected on the basis of a vague contingency. The most accurate scoring possible might never be detected, or if it were, S would not be present at such time to derive the minor benefit of social approval.

Not only would vagueness of contingency run counter to accuracy of scoring but the minor social-approval reinforcer might be inadequate to maintain, in the strength desired, accurate scoring behavior.

Such a scoring procedure is to be contrasted with the following one. The student is told that his errors will be not only of definite benefit in present program revision but that at the end of the lesson his total will be indicated to him. He can see how well he has done in comparison with his classmates, for detailed norms are at his disposal. He can see how his performance compares with theirs in terms of total errors for the lesson as well as whether or not he also tends to miss commonly missed items. Between sessions E checks the accuracy of scoring and comments on it before the next session begins.

This contrast between two hypothetical error scoring procedures has relevance for the study of latencies. Providing latency data in accordance with special instructions may entail behavior which is not followed by reinforcing consequences for the student. Student response may only benefit others, such as the programmer. If it is necessary to depend on the typical student's willingness to cooperate and to depend on a minimum of special reinforcement, his task can be made easy, be guided by clear-cut stimulus events, and when the student behaves in conformity with his instructions this can be indicated promptly to him.

In an initial study it may be desirable to study normal performance, and special instructions may be less of a complication than special instructions plus a reinforcing contingency which favors conforming to these instructions. Furthermore, many measures can be obtained by way of sensing or detecting behavioral processes through the operation of a teaching machine. In such case no special behaviors for the sake of measurement are required of the student.

The Present Study of Latencies

The problem of the present study has already been stated briefly in the first chapter. In this section it will be considered again in somewhat more detail. Five classes of response latencies will be considered: reading, answering, scoring and turning to the next frame of material, reading-plus-answering, and total frame time (reading plus answering plus scoring and turning). As each of these is considered it may be of interest to note some of the issues relevant to current programming and to note some future study made possible by refinements of some of these starting points. In the next chapter the details of the method and procedures of the present study will be given.

Reading. The first easily specified behavior to occur in the presence of a new frame of material is reading. In addition to answering and the early part of scoring, reading is an important component of the operant which terminates with knowledge-of-results reinforcement. The details of this behavior are the details of stimulus control; this behavior more than any other is a function of properties of the stimulus material, the program. Adequate stimulus control of this behavior is essential to reinforcement of the operant which ends with student scoring. (This is not true to the extent that correct responding may occur by chance, and this points up a drawback to the use of items that call for yes or no answers or an otherwise limited class of answer. The fact that the student, rather than the program, supplies both the response class and a member of this class does not keep the probability from being high that correct answers will be given for the wrong reasons, that responding to irrelevant stimulus features will occasionally be reinforced. The high probability of responding correctly due to chance as a result of form of item needs to be vanished or faded as much as do strong prompts.)

An improvement of reading would be an improvement in stimulus control. Since a program often is arranged to make right-answer reinforcement contingent upon adequate reading, an incidental benefit of long-term automated instruction may be improved reading skills.

If this is not true it could become true. The automated instructional procedure, including apparatus and program, could be modified so as to favor the development of reading skills at the same time the student attended to other kinds of program content.

Supplementary Reinforcement. Special reinforcing procedures were developed late in the present study. A stimulus event was made contingent upon latency value.

An electrically-operated counter, which was located on the teaching machine in view of S, provided for bonus points. Increments on this counter were interpreted as indicating unusually good performance, an improvement over previous performance. If necessary, these points could be interpreted as extra credit which applied to S's course grade, or they could be a basis for decreasing some aspects of course work required of him. Various interpretations could be tested with different, small experimental subject groups until these counter increments were found to be reinforcing, i.e., found to increase the probability of the kinds of behaviors which they immediately followed. In a very limited trial, with three Ss, the interpretation of points simply as an indication of improvement seemed adequate.

Such bonus points can be made a function of any kind of latency. Just how the contingency should be specified is likely to be a complex problem in the case of each application. To the extent that latency measures are correlated, a consideration of the present study, other classes of behaviors will change as the one receiving special treatment changes. If faster reading were reinforced, not only might reading get more careless but answering and scoring as well. The kind of stimulus control weakened by reinforcing only in terms of rate is common to all three classes of performance.

Merely identifying desirable behavior in terms of latency for any single class of latency is a challenge in itself. The stability or reliability of behavior--the extent to which it changes from one presentation of a lesson to the next--is a consideration. Using the latencies which were observed on the first presentation of the lesson as a reference point, desired changes could be reinforced as they occurred.

The student could be reinforced for improvement in his reading at the time this improvement occurred during the second presentation of the same material. But, improvement is difficult to specify. It will not always be a matter of reading faster, for often it is necessary to slow down for complex material. It will not always be a matter of the average or poor student doing as the better student, for what is appropriate to the history of one will not always be appropriate to that of the other.

Individual frames of material can be considered. For each frame the median latency of the group of students who responded correctly can be compared with the median latency for the students who made errors. This might suggest direction --whether or not the student should be reinforced for slower or faster reading on a given frame.

Regardless of the details, which can only become specified after the study of such issues, the general nature of the contingency would be that of pacing student reading. Bonus points would be contingent on the student's reading at a rate appropriate to the difficulty of the material. Only reading latencies which fell within the temporal interval appropriate to a given frame of material would result in bonus-point reinforcement.

An important possible modification of such a pacing schedule has to do with varying the size of the interval defining reinforcer availability. Ferster and Skinner (1957), in addition to the concept of pacing discussed above, make available another concept which, similarly, may be applied to automated instruction research, after it has been juggled around a bit. It is the concept of the adjusting schedule.

In an effort to obtain the maximal number of responses per sizable unit of time, Ferster and Skinner demanded less work of a pigeon after some sign of strain was detected, after a long pause. If, e.g., 100 responses were required for each reinforcer and the pigeon paused after being reinforced, when it resumed responding it might have to emit only, say, 75 responses for the next reinforcer. As the previous high rate was recaptured, the number of responses per reinforcer could gradually be increased to the previous value, or an even higher one.

It is suggested that the long pause of a student working at a teaching machine could increase the temporal interval so that, if his next response were an improvement over the pause --had a shorter latency--then even low-grade improvement would be reinforced by a bonus point.

The enlargement of the temporal interval defining reinforcer availability introduces a complication; pausing behavior may be reinforced. The contingency defining the non-adjusting portion of the interval should be such as to favor this class of latency. Otherwise, the pause, instead of initiating an adjustment which led to recaptured performance, would become the response which increased in frequency.

Since the pause merely changes the contingency, two responses are required for each pause-response reinforcer. Prior to enlargement of interval, the basic contingency should make available one reinforcer for more than 50 per cent of the responses. Such a contingency would not favor pausing.

It should be noted, in addition, that student performance would tend to be maintained by the regular knowledge-of-results reinforcement. The additional source of reinforcement from bonus points for improvement following strain merely complement the regular source. Pausing is not only not favored by the adjusting contingency, it delays another kind of reinforcement.

Answering. As an important component in the program-working operant, answering immediately precedes the terminal component of scoring. It does not immediately follow reading; between reading and answering there is the important covert behavior of "thinking." If this intervening component could be adequately measured, such a measure could be of interest as a new operation which helped define degree of composition in the composed answer. In the sense of such a new definition, an answer that required 50 seconds of composition, "thinking," or covert responding time could be considered more composed than one that required, say, 5 seconds.

Composing time could be used in relation to bonus points to meet an objection noted by Skinner (1958). Some students commented that once they got started on a lesson they did not feel like they had time to think. Bonus points could be pitted against reinforcement by progress toward the completion of the lesson. However, it is not clear how much of this anxiousness to finish was due to specific features of the machine used by Skinner, such as numbered frames on a disk.

If there is a general problem of undesirable haste being favored by the typical teaching machine, the place to oppose this haste may not be the total frame time but, rather, something as specific as composition time. However, the pacing of reading could ameliorate this difficulty.

Overt answering begins where covert answering leaves off. The fact that this may happen several times on a single frame of material, that the student may go back and forth from covert to overt answering, complicates a separation of the two. When the student is asked to indicate when he has his answer in mind ready to write down, overt answering times are much longer than are required merely for writing his answers.

A further complication in the present study was the extent to which "thinking" time could be assimilated to either reading or overt answering. Some of the most obvious approaches to this problem would involve controlled visual access and controlled access to the answer-space. What was neither reading nor answering would be the covert component. These special restraints or complications of normal use can wait until some data has accrued in the case of less radical machine modification. Reading and answering measures may eventually be studied when they are obtained under conditions affording such refinement, and this may be necessary to any of the future studies discussed in this chapter.

Probably the most important characteristic of the overt answer is its topography--whether it is a correct answer. But correct answers may differ among themselves as to a time requirement (as may incorrect ones). Latency is another dimension. It may, e.g., indicate something for the response that has no topography. For a student to spend two minutes on an item he finally leaves blank may be quite different than for him to give up trying to answer after only five seconds. If the programmer were trying to program some control of persistence in problem-solving, latencies associated with blank answer-tape could be an important datum. (While this datum should be scarce in an easy-item program, even here an end-of-lesson review test may be an adequate source.)

In view of the work of Glaser, Homme, and Evans (1959) the relation between overt and covert responding may become important as a programming issue. If students can really learn as much through covert as overt responding, their time and effort may be saved. But what feedback would there be to the programmer? Of course, the program can already have been perfected with overt responses. However, the programmer or investigator would probably like some assurance that the student really was responding in the covert instance. Glaser, Homme, and Evans report that they did not find significant differences "between Ss who were instructed to make overt written responses to each step in the learning sequence and those who were instructed not to make written responses" (Lumsdaine & Glaser, 1960, p. 451).

Reading-plus-answering. Latencies for this category are for the complete operant if one excludes some fuzziness of termination at the time of scoring. This unit and its components in the overt case can be studied in relation to the covert one.

As a measure its reliability is favored by excellent S^D control, a stable contingency, and knowledge-of-results reinforcement. It is also a measure which can be obtained without special instructions or special procedure; it can be obtained through switch action occasioned by the internal mechanism of a teaching machine during normal operation.

As a measure it is superior to total-frame-time in that the time spent pondering an incorrect answer and the somewhat artifactual small component of time due to operation of an error scoring mechanism are excluded. It reflects a latency obtained before an error is recognized as an error by the student.

Scoring-and-turning. As soon as the student had finished reading and answering he could immediately compare his answer with that of the program. As the shutter-mechanism for this comparison was operated, answering time ended. Scoring is the behavior most intimately related to the reinforcing event. The latency component which differs in the case of right and wrong answers, if there is a difference, lies within the scoring component of this scoring-and-turning measure.

Turning is the behavior of advancing the program to the next frame of material. It is an instrumental action which produces the S^D for the program-working operant. (A case could be defended that turning is actually the beginning of this operant.) Mixed with, and lying between, the two components, scoring and turning, is more "thinking" time, as the student ponders his answer before, and as he turns. These three components all contribute in unknown amount from each to yield the measure called scoring-and-turning.

Some consideration of scoring behavior may help bring out one of the basic limitations of essentially descriptive studies such as the present study. It is a correlational approach which, at times, seems to offer information which often later turns out to be highly ambiguous. Experimental treatments under controlled conditions are a necessary complement. It will require some development of the issue which seems to illustrate this point.

While providing difficult material which is conducive to error may be undesirable programming, student response at the time of scoring may determine whether or not there is any benefit from his errors.

Ideally the student reads a question and answers it correctly. However, some items may be missed. If the student repeats a lesson and gets previously missed items correct on the second presentation, the question context of the material is no longer as exclusively the discriminative stimulus. The student's recall of seeing the answer blends with the question as a contribution to his answering the question correctly on the second attempt. The student did not strictly "arrive at the answer"; he remembers seeing it. Even worse is the case in which items are repeatedly missed.

Meyer (1960) reports a lack of success in one attempt to deal with this problem. She found that a group of students who simply made an "X" on wrong answers did not differ significantly from a group who later went over pages containing errors and wrote another answer. While it might seem that the second group would have additional opportunity to be correct, and thus be reinforced, this did not result. Instead, five students of the second group failed to score an omission of answer whereas none of the first made this scoring error. Faulty scoring developed more readily than the behavior of learning to give correct answers when provided a second opportunity.

A special bonus point contingency might be applied in order to develop scoring behavior that decreased the probability of repeated error. Here is a possible interpretation of error repetition. An answer which the student gives which does not match that of the program is probably an important component of an aversive stimulus situation. The mismatch of answers probably is most aversive initially. As escape and emotional behaviors subside, which are incompatible with the kind of stimulus control required for careful reading, the student becomes more capable of covert responding such as reading the program answer correctly to himself. In addition, he may reflect a bit, relating some features of the correct answer to previous instances of response to the program. The closest this covert, reflective behavior can presently come to being amenable to objective study would be in terms of a loose relationship such as this: If it occurs it takes time; unusual amounts of scoring time indicate unusual amounts of reflecting.

Of course this is not completely true. But, longer pondering time could indicate greater opportunity for incompatible emotional behaviors to subside as well as some such covert analysis and relating of subject matter issues.

At first glance, the data of the present study should bear on this issue. Errors which were repeated on the second presentation of a lesson could be contrasted with those made on the first but not repeated on the second; the scoring-and-turning latencies for the first presentation associated with these two classes of error (those made only on the first presentation and those made on both the first and second presentation) could be compared.

The fault of this comparison becomes clear when it is realized that items which are more difficult are likely to require more scoring time as well as be more likely to be missed again. If difficulty were held constant and scoring time manipulated by bonus point supplementary reinforcement, some conclusion regarding this issue might be possible.

Total frame time. From the consideration of small components of this larger unit, which often indicated a need for further refinement, total frame time seems most useful as an indicator that within a given frame of material somewhere there is or is not faulty programming. It is a starting point much as error is a starting point. The content of an error helps the programmer in much the same way that components of this overall frame latency may help him, as it becomes possible to interpret them.

Long latencies may be expected to indicate item difficulty to the extent that they can not otherwise be accounted for in terms of the amount of material to be read or the number and size of answer units.

Special problems involved in the use of supplementary reinforcement. The possibility that stimulus control may be weakened as a result of reinforcing in terms of rate has already been mentioned. Careless reading may result from reinforcing fast reading. The student may fail to see or score his errors.

There are several issues. Reinforcement needs to be given promptly--at the time of fast reading, not at the end of the frame after scoring. But it needs to be contingent on a topography which indicates adequate stimulus control, and this can not be known at the time of fast or slow reading (whichever is being reinforced for that frame). If reinforcement can be given in a way to favor some discrimination of what kind of behavior is responsible for reinforcement there should be some gain.

For such considerations, it seemed desirable to have a light flash in the event that a reading latency fell within the desired time interval. The rate of flash was a function of the rate of reading being reinforced. After the flashing lights the bonus point could be given or not at the counter depending on whether or not an error had been made. The S always experienced the flash even though scoring his answer as incorrect prevented bonus point delivery. The flash came immediately after reading rate had been established; but the bonus point, if S did not indicate error, came later: not during scoring but after turning had been completed.

Temporal relations should favor bonus point control of rate of reading more than the act of overlooking errors even though the latter came closer in time to the reinforcing event. The flashing light was to become a conditioned reinforcer on the basis of intermittent pairing with points. And, although weak, it would be a timely stimulus change. The work of Zimmerman (1957) suggests that greater resistance to extinction of conditioned reinforcers may be gained by way of intermittent pairing. Minimal reinforcement at the optimal time, which is provided by a stimulus that can be depended upon to maintain its reinforcing capability, may favor specificity of behavioral change.

More may be needed to maintain scoring accuracy. Some bonus points may need to be given in those cases in which S makes an error yet spends an unusual amount of time scoring and turning (and, it is hoped, pondering). However, it would probably be a better strategy to make accurate scoring of value to S through allowing him the use of norms. Comparisons would only be meaningful if scoring were accurate. A teaching machine which scored would even be better as a means of assuring rate manipulation without reduction of stimulus control.

Summary

This chapter has considered the measurement of latencies in general and in the present study in particular. In both cases there has been emphasis on the nature of the behavioral unit responsible for the measure. Measures will be a function of the sharpness of stimulus control, of the discriminative stimulus, of the stability and simplicity of the contingency which defines the conditions behaviors must satisfy in order to be reinforced, and the nature of the reinforcer.

Many measures are possible through the detection of teaching machine operations such as positioning new materials, writing an answer, operating a shutter for scoring, operating a scoring mechanism to indicate an error, etc. More are possible on the basis of special instructions to S asking him to indicate when his response to some stimulus event is complete (such as his reading of a new frame of material) or when his behavior provides a stimulus event necessary for other behaviors (when he has an answer in mind that he next writes down).

Latency measures may help describe behavior which is largely covert. For this reason, they may prove useful in the study of covert compared with overt answering. It is one thing to compare the covert answer with the overt answer; it is another to compare the overt answer with some other covert process.

The five categories of latency considered in the present study were considered in relation to other possible future studies. In addition to serving to describe, they make possible new kinds of control. The student who loses interest and pauses only to resume more normal machine operation on the next frame earns a bonus point. In addition to improvement during a session, bonus points may be given for performance on a second presentation which is an improvement over that of the first. These matters received only slight study, although there has been extensive planning of supplementary reinforcement procedures.

For example, reading could be paced in a manner appropriate to the difficulty of material. Eventually, difficulty might come to be more of a controlling factor as S came to be reinforced for both faster and slower reading, depending on characteristics of the material.

Slow, careful reading of a frame after scoring an error may deserve reinforcement. The prevention of error repetition seems to be second, as a programming goal, only to placing items at the desirable level of difficulty.

Change in any program-working behavior may be associated with change in another. Scoring accuracy deserves consideration after any special procedure is introduced.

Chapter IV

Method

Subjects

The Ss were 12 males and 4 females obtained from six remedial English classes. Four of the males came from a class that completed only the early portion of a remedial English program during the fall. These students continued to work on new lessons as these became available. All other Ss were currently enrolled in one of five sections of remedial English during the spring and served in the experiment as part of their regular class work. The data of one female S was of special interest; she was a foreign student who seemed to have much difficulty understanding English. Both written and spoken explanations were of limited success at first in matters concerning experimental appointments and procedure.

These Ss differed in the extent to which they participated. The range was from two one-hour sessions per week over a period of six weeks to only two sessions, one of which was given over entirely to training and instructions.

In addition to this small sample of Ss who participated directly by working on programmed materials on a modified teaching machine, there were 129 Ss whose data were made available as a result of regular classroom activity: work on programmed material in booklet form.

Apparatus

The programmed material was prepared by duplicator, and the paper used was cardstock (with the fanfold creases running parallel to the grain of the paper). A modified Rheem Califone Didak 501 teaching machine (a preproduction model) was used in this study.

One feature of this particular model was of importance to this study. Production models are not expected to include this feature: a brake controlled by holes punched in the program. These holes allow a pin to come through and apply the brake so that programmed material is properly positioned in the viewing window. This feature provides for a standard presentation of stimulus material with a minimum of attention from the student operator.

Action of this brake-mechanism pin operated a microswitch that was fitted into the machine as a part of the modification for the sake of this study. As S brought a new frame of material into view, the microswitch closed, thus causing pulses, generated one per second by a Hunter timer, to be routed into the first of three channels of an Elmeg, ZDG III 2/1, print-out counter.

The action of the brake and the sound of relay action several feet behind S were reliable indicators to S that material was available for him to read. Later, lights were added that came on at this time and augmented the complex of events which defined the beginning of reading time. S was instructed and trained not to read before brake action occurred.

As S wrote his answer, the pressure of his pencil operated a second microswitch. Installed as part of the internal mechanism of the machine, this microswitch was connected by a thin rod to a metal plate located under the answer-tape upon which S wrote. The operation of this switch routed the one-per-second pulses from the first channel of the counter (reading time) to the second channel which began accumulating these pulses for answer time.

Since it was difficult to maintain an optimal range of microswitch sensitivity for all Ss, several strategies were adopted in an effort to separate reading time from answering time. Ss were asked to write the number of parts required by their answer, as writing numbers seemed to involve greater pressure than other writing (and Ss did sometime fail to respond to all parts of a frame of material). S was provided a hard lead pencil; the metal plate under the answer-tape was made slightly rough in an effort to translate some of the horizontal movement of the pencil into downward pressure on the rod to the microswitch.

The best solution to this problem seemed to result from having more than one microswitch acted upon by depression of the metal plate. Wired in parallel, closure of either one served to activate a relay which remained latched upon operation. Once pulses were routed into the answer-time channel, S did not have to hold the answer-tape plate down for pulses to continue to accumulate in this second channel. The rod of this additional switch was located directly beneath the spot where S seemed most often to begin writing.

The switch itself was extremely easily operated yet not likely to be operated by S during non-writing machine operation. A greater excursion was required for switch closure by the introduction of a lever between the switch and the rod. The rod acted on this lever which, in turn, acted on the switch. The leverage which made the switch easier to operate also increased the plate-excursion, making accidental activation due to jostling of the teaching machine less likely.

During the time the pin was through the punched hole in the program, pulses were accumulated in either the first, reading-time, channel or the second channel as answer time. The total count for each channel was stored for later printout.

The separation of reading time from answering time was accomplished by sensing the first instance of S's writing. S was instructed not to begin writing until he had all parts of the answer in mind for the entire frame. As was mentioned earlier, S could "think" while reading, while answering, or do some thinking in the course of both activities.

The time accumulated, from the moment a frame of material was positioned in the viewing window until S operated a shutter revealing the correct answer for scoring purposes, was available as a read-plus-answer time which could be obtained by adding together the count accumulated in the first and second channels, somewhat gratuitously (due to the complication of "thinking") referred to as reading and answering times.

To keep these measures meaningful to even this extent, it was necessary to add lights that were turned out by S when he wrote his answer with sufficient pressure to provide the intended count. These lights were especially needed during the early development of a means for measuring answer time. Then it was especially difficult for S always to conform, and to know that he had conformed, to his instructions of bearing down firmly as he began to write his answer. And, as was noted earlier, the onset of these lights was of some value as a stimulus change which indicated to S that he should then, and only then, begin reading the material which he had just placed in view in the window.

Scoring and turning to the next frame was also timed. The fact that the brake-mechanism pin moved down through the hole in the program when S opened the answer-shutter to score his answer helped make this possible. Once the pin was down it remained down, being kept down by the paper, as S advanced the program, until the next punched hole was encountered. Cardstock-weight paper not only resisted wear but favored crisp action as the pin came through the firm-bordered hole, yet flexibility was gained by taking paper grain into account.

When this pin was through the hole, as was mentioned previously, the first microswitch was closed. This operated a relay and provided a circuit through a set of contacts for the definition of reading time. Other contacts on this relay provided the circuit for scoring-and-turning time when the microswitch was in the open position.

The pulses of this circuit entered the third channel of the print-out counter and were stored there to be printed out with the count of the other two channels. The simultaneous print-out of all three channels occurred as scoring-and-turning ended, as a result of the pin encountering the punched hole associated with the next frame of material.

As the pin came through the newly encountered hole, a print-out of the pulses accumulated in all three channels occurred, appearing as three separate totals printed on one line of a paper record. A re-set followed so that all channels started at zero at the beginning of the next frame. During this time of print-out and re-set, one line of recording paper was also advanced through the print-out counter.

The noise from this action coincided with the onset of the lights and the brake action of the teaching machine and signalled to S the availability of a new frame of material. The action of print-out, paper advance, and re-set, although occurring after the brake action and the turning on of the lights, did not introduce any sizable constant error in reading time because it occurred at the time S was turning his head and focusing his eyes on new material. Approximately 400 msec. were required for these actions of the counter. Many Ss seemed to have just oriented themselves to the new material as new pulses began to accumulate in the reading time channel.

Later, when bonus points were introduced and given a very limited test, if S earned a point its appearance at the counter on the teaching machine also came at the end of scoring-and-turning time.

Materials

The need for samples of material, as well as subjects, resulted in some compromise, for practical reasons, in the size of each. Three units from an eight-unit remedial English program--units II, IV, and VII-- were the source of 5, 2, and 6 lessons respectively. The shortest of these lessons contained 50 frames of material, the longest, 100. After finding that the latter lesson, in the case of two Ss, required longer than the usual hour, it was made into two lessons.

Although the program usually involved a composed answer, some frames required yes or no answers as well as the occasional selection of a correct alternative from three or more possibilities. Sometimes the choice was indicated by the letter or number associated with the item chosen; sometimes the full content of the answer was written.

The program emphasized concepts of traditional grammar; all Ss had completed at least 13 lessons involving parts of speech before contributing data to the present study. The materials studied under experimental conditions treated sentence structure and punctuation.

Unit II had been revised two or three times on the basis of classroom testing. None of Unit IV had been revised. Part of VII had been revised on the basis of testing with four students: lessons 1, 2, and 3.

Procedure

Some of the first students to finish Unit I in booklet form as part of their regular classroom work were given an opportunity to continue their work on a modified teaching machine. Later, average and some of the slowest students were enlisted in the modified-machine group. S could work at the same time of day as he had been in class or any other time that the machine was not in use. Each hour of work could be substituted for an hour of regular classroom attendance.

Students were told that they would help test a teaching machine so that more could be purchased if it proved satisfactory. In addition, they would help E discover parts of the program which took either too much or too little time. S was not given any detail as to how timing was done. It was emphasized that it was the program and the machine that was of interest; S's behavior was only a means to this end.

S was asked not to make any comments to E during his work on a lesson. If he had any problem with the material he could make a checkmark on his answer-tape, and these points of difficulty could be considered at the end of the session. The time spent on various parts of the task was of interest.

Therefore, S would need to make his own decisions about how to proceed. If he had to take time to puzzle things out, that would be discovered in the record.

Aside from such special instructions S was to do his regular course work on the machine. His program would continue to be that of the regular classroom; he would have the same exams and outside composition assignments. Occasionally he would repeat a lesson for either his own benefit or that of a special experimental purpose. But, care would be taken that little was asked of him that was of value only to the experiment.

During the first session S was instructed in the use of the machine. He was told to apply strict scoring standards as this would help point out imperfections in the program. His error mark should indicate any disagreement between his answer and that given by the program even if it were obvious that the program was at fault.

S was asked not to read material as he turned the knobs which brought it into view. He was asked to turn his head, avoiding a view of the window, until he felt the brake go on and heard the noise of equipment action. Then he should begin reading immediately at his own normal rate.

Only when S had completed reading a frame of material and had all answers in mind was he to begin writing his answers. As he began writing he was to bear down firmly on his pencil and turn out the lights, lights which tended to come into view more as S looked toward the answer-space. (S was expected to become more aware of them just before making the response which could turn them out, just before writing his answer.)

Since no attempt was made to obtain meaningful data on the first session, E discussed procedure with S as he found it necessary. S was reminded that he had failed to put out the light at the desired point in the sequence if this became necessary. He was asked to turn his head away from the window if it seemed he might be reading before his start signal. E explained how S was to operate the teaching machine in matters such as scoring errors (e.g., with the force needed to operate the error counter, which would come into view at the end of the lesson).

The material which S worked on during this first session was a second presentation of a very simple lesson, the first of the program. Most Ss had already worked this in booklet form. S was encouraged to ask questions for he would be expected to conform to somewhat standard conditions on the next session. Then he would have to avoid any questions during the session, work on his lesson without interruption until its completion, but do his work at a normal rate, one that was comfortable to him.

An important part of the procedure concerns apparatus. The pulse-per-second performance of the Hunter timer and print-out counter was considered in a calibration study before work with students began. Error was slight, only a little greater than that of the watch and clock combination used in a few tests in which pulses were accumulated for an hour or so.

During the study, reading time measures were occasionally lost due to a partial impression of one digit on the printed record. These and other kinds of recording failures, as well as measures lost due to atypical experimental circumstances, contribute to some rather complex forms of bias.

There were many independent reasons for the loss of measures. Equipment failures during the course of the study involved a defective relay, a Hunter timer failure (apparently due to house current fluctuation at the simultaneous start of many airconditioners), a malfunction of the print-out feature of the counter, a blown fuse, and several mechanical failings of the teaching machine.

Some bias results from the fact that later in the day there was slightly more tendency for apparatus failures. There was less time during the course of the day for repairs than there was at the end. Also, everything used by the first S of the day could be more thoroughly tested ahead of time.

Another source of bias stems from the fact that there are large individual differences in the use of mechanical devices. Explanations and suggestions from E were insufficient to help all Ss with their work at the machine. And the way in which the machine was used sometimes determined whether or not data were obtained or were meaningful. Mechanical features of the device used tended to exclude some data from Ss who would have been better able to contribute had the controls been more fully automatic, say, push-button.

Both during the experimental session and during the analysis of data a variety of safeguards were adopted in order to detect faulty data. E was in a small room that permitted him to watch the equipment in an adjoining room as well as watch S as he worked at the machine (which was in another small room beyond that containing the equipment).

Lights at the print-out counter flashed to indicate which channel was receiving pulses. The flash as well as the sound of relay action indicated whether or not the timing system was functioning normally. E could advance the print-out record at the time of any irregularity and indicate by a blank line that the next contained the irregular values. Sometimes only one measure was lost, as when S asked a question during scoring-and-turning time. Minor irregularities were ignored, such as S dropping his pencil.

Each hole of the program was an occasion for microswitch action and the print-out of one line of figures. Each printed lesson should have the same number of lines as there were punched holes for the lesson. Machine failures which required opening the machine and mending the program or adjustment of the answer-tape complicated the record. Often there were useable portions of the record in the case of one or two irregularities of this kind.

Several kinds of criteria had to agree for irregular data to be included in the analysis. The sequence of values had to have either a known start or end value that related lesson to record. Where there were review lessons, the three zero value answer times had to agree with the three instructional frames as far as sequence was concerned. A long, short, and long reading time should match the portion of the program that had two full windows of material, one before and one after an unusually brief frame, etc.

Data obtained from the regular classroom use of the materials in booklet form indicated implausably low error rates in the case of student self-scoring. On some frames, E-scored records of Ss who worked on the modified machine revealed an error rate ten times as high as that suggested by classroom student-scored records.

Classroom review tests, which were scored by a clerical assistant, had about the same error rate as those scored by E for students doing their work on the modified machine.

As students in the classroom completed the review test at the end of certain lessons, they handed in their papers in order to be given temporary access to the scoring key. Their paper was not returned to them at this time.

Chapter V

Results

Subject and Lesson Characteristics

A total of 16 subjects participated in this study. In many of the following tables, results are presented for Ss according to an identifying number. The four female Ss are numbers 2, 3, 12, and 16. Subject #12 is a foreign student who had unusual difficulty with English. Of the remaining Ss (all male), numbers 4, 5, 8, and 11 were continuing work begun the previous semester.

Thirteen lessons were considered. Lessons are numbered in arabic, units of material in roman numerals. Three units were the source of these thirteen lessons. Materials from all lessons are basic to several of the major findings. However, some lessons from Unit II received special consideration for three reasons: (1) A greater number of Ss completed a part of Unit II. (2) It was the only unit with reviews. Four of the five lessons were followed by review tests: 1, 2, 3, and 5. (3) Unit II had been revised on the basis of classroom use. Revision may have made this material more similar to other programs under development or in use than would be the case with first-version material. (The per cent error of Lesson 3 of Unit II was approximately 10. Therefore, it is likely that this material was slightly less difficult, for the students of the experimental group of this study, than was the psychology program considered by Holland for the psychology students who reacted to the second version. Holland reported an 11 per cent error for his revised material.)

Sample frames of material from three lessons of Unit II are given in Appendix A. Included is information concerning error and latency as related to these frames.

All of the review test material of Unit II is given in Appendix B. Classroom error rate for each frame is given, based on the scoring of a clerical assistant rather than the student. Since these tests are based entirely on the programmed material of the lesson, from the review it is possible to infer the nature of the lesson material--especially when such inference is aided by the sample frames of lessons given in Appendix A.

None of the four Ss that were continuing work begun the previous semester worked on Unit II during this study. These continuing Ss were the only Ss to work on Lessons 5, 6, and 7 of Unit VII and Lessons 1 and 2 of Unit IV.

Data Reduction

Table 1 is a record for one subject on the first presentation of a lesson. Values in three columns of Table 1 were obtained directly from the print-out counter record: Read, Answer, and (Score-&-Turn). The other two kinds of latency values were derived from these, as the labels suggest: (Read-plus-Answer) and Total Frame, (Read-plus-Answer) plus (Score-&-Turn). Table 1 illustrates raw data and a step in which two kinds of latency measure are derived on the basis of these two

Table 1

Response Latencies for Subject #6 for Lesson 3-II on First Presentation

(Individual frames of this lesson for each of five latency categories; data obtained on first 62 frames of this 66 frame lesson)

Frame	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1	4	6	10	3	13
2	6	6	12	3	15
3	7	10	17	3	20
4	3	8	11	3	14
5	8	12	20	3	23
6	18a	33a	51a	12a	63a
7	6	10	16	2	18
8	6	5	11	4	15
9	18a	0a	18a	11a	29a
10	7	5	12	4	16
11	7	7	14	3	17
12	6	6	12	3	15
13	4	12	16	4	20
14	6	4	10	3	13
15	12	4	16	2	18
16	9	5	14	4	18
17	11	4	15	3	18
18	10	12	22	4	26
19	10b	6b	16b	8b	24b
20	20	3	23	3	26
21	4	2	6	2	8
22	5	3	8	2	10
23	18	5	23	4	27
24	19	6	25	4	29
25	35	12	47	4	51
26	12	9	21	4	25
27	28a	0a	28a	4a	32a
28	38	4	42	3	45
29	44a	4a	48a	7a	55a
30	17	5	22	3	25
31	11c	4c	15c	18c	33c

(Table continued on next page)

a Student-scored error.

b Long Score-&-Turn time for frame on which S seemed to see, but not score, error.

c Faulty frame, program repeated answer word. S made special mark on answer-tape.

Table 1 (Continued)

Frame	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
32	7	4	11	3	14
33	14a	18a	32a	6a	38a
34	16	7	23	3	26
35	24	5	29	2	31
36	13	4	17	3	20
37	26	13	39	4	43
38	6	10	16	3	19
39	18	12	30	4	34
40	19	7	26	2	28
41	20	10	30	2	32
42	14	18	32	7	39
43	7	34	41	9	50
44	6	17	23	6	29
45	18	10	28	11	39
46	12	10	22	3	25
47	12	4	16	5	21
48	9	11	20	3	23
49	29	4	33	5	38
50	5	9	14	4	18
51	11	10	21	4	25
52	18a	9a	27a	5a	32a
53	5	5	10	3	13
54	3	3	6	3	9
55	7	7	14	3	17
56	7	5	12	3	15
57	15	4	19	2	21
58	31	44	75	5	80
59	18	15	33	2	35
60	8	6	14	2	16
61	12	9	21	4	25
62	22	4	26	2	28
Average:	13.56	8.71	22.27	4.24	26.52
Range:	3-44	0-44	6-75	2-10	8-80
EAA/EBA:	6/0	3/3	5/1	5/1	6/0

a Student-scored error.

combinations of raw data. Since more than 100 pages of such tables would be required to give this kind of detail, such records are not presented.

The performance of Subject #6 on Lesson 3 of Unit II is offered primarily as an example of a record. A secondary consideration was that this particular S completed several lessons. Consequently, Table 1 may be related to other tables which present findings involving this S. The performance of Subject #6, during the course of this study, may illustrate several data-reducing procedures as well as be a basis for some of the findings.

Similarly, Lesson 3 of Unit II was completed by 7 Ss and can later be related to other tables. It is similar to other lessons completed by as many Ss in terms of subject matter, an end-of-lesson review test, and stage of revision. Also, as was pointed out, the error rate of this lesson suggests that, as far as this kind of difficulty is concerned, it seems to be slightly easier than the well-known Holland-Skinner program.

The values in the body of Table 1 are times, in seconds, spent on a single frame of material for the kinds of activities given at the top of each of the five columns. In one sense these are response times, The time spent reading, e.g., is the time required for various reading responses to occur. In another sense, these are latencies. The operant of reading to completion, when timed from S^D (new frame of material) to end, yields an operant latency.

In all tables, time is given in seconds. Throughout, no effort has been made to make a distinction between response time and latency.

To summarize the various kinds of latency or frame times, five for each of the frames considered (e.g., such as the 62 frames of Table 1), averages were computed as a first step in data reduction. For the number of cases typically involved, averages were easier to obtain than medians (especially if the same record was considered for a Pearson product-moment correlation problem). Ranges were even more a matter of convenience, as compared with standard deviations or variances, as an index of variability for a set of values.

Records for the various Ss, for each lesson they completed, were summarized (in terms of each kind of latency) by an average and a range. In addition, the errors made on each lesson were classified in terms of associated latency value. Errors were classified above or below average with respect to each kind of latency average, Read, Answer, etc. Thus, for most records, the errors for a lesson were successively classified in terms of five kinds of latency average. Table 1 is an example showing this for one S on one lesson.

At the end of Table 1, the average time spent reading a frame of material is given as 13.56 seconds. The least time spent reading a frame was 3 seconds, the most, 44. Subject #6 made 6 errors on this lesson portion of Lesson 3 of Unit II;

above average reading times (14 seconds or more) were associated with all frames on which these errors occurred. This S spent 8.71 seconds answering the average frame. On at least one frame he either failed to give an answer or failed to operate the answer-switch (judging from the zero lower limit of the range). On one, he spent 44 seconds answering, supposedly after he had completed reading. Here, in the case of answering, half of his errors were associated with above average answer times, half with below. (However, inspection of individual frames reveals that two errors are associated with zero latencies; S did not answer. Instead of zero time, S spent infinitely long in answering these two frames, one might argue. Thus, two errors should be excluded rather than classified as below average, The problem here is that Answer refers to overt answer; zero values need to be taken into account in specifying the average time per frame spent in overt answering. Later, in an analysis where this issue seems most important, results will be reported both with zero values included and with them excluded.)

Since, in Table 1, Read, Answer, and (Read-plus-Answer) are all based on the same number of observations, the average for Read, when added to the average for Answer, yields the same value as the average for (Read-plus-Answer).

As far as the lesson as a whole is concerned, there are four missing observations. A convenient way to express this fact for the sake of description or comparison with other lessons is in terms of this fraction: 62/66. Observations were obtained on 62 frames of a possible 66. One way or another, data were lost during recording (or could not be interpreted) from four frames of this lesson.

In addition to this statement of frames providing data relative to the total number for that lesson, all records which provided data from (1) a first presentation, (2) a second presentation, and (3) a review test were reduced to (i) average values in the case of each kind of latency, (ii) ranges for each, and (iii) classifications of errors in terms of whether or not associated latencies fell above or below each kind of average.

Such intermediate steps in data reduction are given in Appendix C by Tables 16, 17, and 18.

Of interest in the study of latency data is the extent to which the various kinds of values vary from lesson-to-lesson. Table 2 is given to illustrate the performance of one S on eight different lessons. Individual differences were held constant while lessons and such factors as occasion vary. Of course, this means that differences between lessons could be due to almost anything--except individual differences.

One change which seems most likely to be due to a characteristic of the material is the proportionately shorter answering times on the three lessons of Unit VII. The unusually low Score-&-Turn average suggests a possible surge of interest in the first lesson dealing with punctuation, although this could be due to a change in response mode. Emphasis on marks of punctuation, rather than words or phrases, may help account for the change. Thus, the time required for a careful check of spelling may be saved at the same time that simple frames of material calling for marks-of-punctuation responses are introduced. Yes- or no-answers and other forms of two-alternative items may have been a factor here, also.

However Table 2 may be interpreted, the greatest change in latency associated with changes in material is in terms of the relation of Read to Answer latencies.

Table 2 illustrates another step in data reduction. Since only a few lessons are involved in the case of each S, it becomes convenient, and desirable, to obtain the median value of his lesson averages. Medians, less sensitive to a few extreme values, are preferable (when not precluded by computational difficulty) to averages in the context of the present study where early recording problems may have led to the inclusion of a few faulty measures, e.g., Answer time.

Table 2 shows the procedure, in the case of one S, that was followed in the case of each S. The median value for each kind of average latency (where the number of averages contributing to this median was the number of lessons) was obtained as a way of indicating typical performance for each S. Such medians over lessons, as is illustrated for Subject #6 by Table 2, summarize the total participation of S. These medians were obtained separately in the case of first and second presentations of a lesson. The median of averages for lessons obtained from a first presentation of a lesson is not

Table 2

Response Latencies for Subject #6 on First Presentation of Eight Lessons

(Average frame of each lesson for five latency categories)

Lesson	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1-II	9.65	8.07	17.72	4.30	22.02
2-II	12.78	12.90	25.67	4.33	30.00
3-II	13.56	8.71	22.27	4.24	26.52
4-II	17.67	7.69	25.36	3.64	29.00
5-II	10.15	10.15	20.31	3.98	24.29
1-VII	10.59	4.06	14.65	2.85	17.50
2-VII	11.28	4.08	15.36	3.38	18.74
3-VII	12.42	4.33	15.39	3.78	18.86
Med.	11.85	7.88	19.02	3.88	23.16

affected by data obtained on a second presentation.

Table 2 suggests that it was "typical" for S #6 to spend 11.85 seconds reading the typical frame of material, 7.88 overtly answering, and 19.02 on the combination (which is not the sum of these two since establishing the value for Read did not set the value of the median for Answer nor for Read-plus-Answer). It was characteristic of this S to spend 3.88 seconds scoring his answer and turning to a new frame of material. He usually spent 23.16 seconds on a frame as a whole, carrying out each of these activities.

Table 1 can be related to Table 2 by noting that the row of average latency values of 3-II consists of the values given as Average at the end of Table 1.

The procedure applied in Table 2 in the case of Subject #6 was followed with all Ss, and the results appear in Table 3. The entire participation of all 16 Ss on the first presentation of lessons is given by such medians based on lesson averages for each of the five kinds of latency. The median values of Table 2 appear as the row for S #6 in Table 3.

Table 3 provides latency values which are characteristic of the 16 Ss who participated in this study. These values are characteristic of their entire participation on the first presentation of lessons. The median of these medians gives the latency values characteristic of a typical subject on a typical lesson, as far as this study is concerned, for each of the five kinds of latency.

Differences between values in each of the columns is

Table 3

Typical Response Latencies for Each Subject on First Presentation

(Median based on average frame of a lesson for all lessons completed by S; five categories of latency and 16 Ss)

S#	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1	14.74	13.86	28.60	5.44	34.05
2	12.84	8.43	21.28	4.34	25.61
3	15.29	10.24	22.77	3.09	26.45
4	12.01	11.03	22.44	6.74	28.96
5	14.88	11.78	25.38	5.38	30.24
6	11.85	7.88	19.02	3.88	23.16
7	13.47	13.49	26.96	5.23	32.18
8	16.16	8.83	24.99	4.54	29.52
9	15.16	8.11	23.27	3.38	28.94
10	14.67	11.35	26.01	6.45	31.79
11	13.76	9.98	22.36	3.94	26.93
12	20.87	15.22	36.04	9.38	45.52
13	10.78	18.10	28.87	7.02	35.89
14	23.85	10.58	34.42	5.15	39.58
15	12.70	7.27	19.39	4.32	23.73
16	9.04	14.25	23.29	4.60	27.88
Med.	14.22	10.81	24.14	4.88	29.24

largely a result of each row's being based on a different S, although the various Ss completed different lessons. Table 3 points up the kind of differences a general procedure involving latencies would have to accommodate if such generality went so far as to ignore both differences due to materials and individual differences, in establishing bonus point contingencies, e.g.

Table 4 gives similar information for a second presentation of a lesson. More detail concerning the number of lessons or the number of frames of each lesson providing latency data may be obtained from the tables of Appendix C.

It should be noted that whereas only three of the sets of medians given in Table 3 are medians only in the sense of being a median based on one case, four (and thus, proportionately more) in Table 7 are medians in this questionable sense.

Latencies for Lessons

Table 5 shows the differences among latencies for the seven Ss who worked the same lesson. It is for a lesson what Table 2 was for a subject. In Table 5, the differences between rows are due to differences between individuals (complicated, of course, by factors that go with differences in occasion). Material is held constant.

Since Lesson 3 of Unit II has been important in several tables, it may be useful to future investigators to know some of the characteristics of this material.

Although the error rate given in the case of the seven Ss in the experimental group was only slightly below that for the

Table 4

Typical Response Latencies for Each Subject on Second Presentation

(Median based on average frame of a lesson for all lessons completed by S; five categories of latency and 8 Ss)

S#	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
3	8.19	9.15	16.57	2.71	20.06
5	8.58	11.26	19.84	4.70	24.39
6	9.80	6.16	15.95	2.97	18.92
7	8.65	8.65	16.32	3.65	21,26
8	15.83	8.36	24.16	5.04	29.12
10	11.54	10.65	22.19	4.88	26.12
11	10.31	9.19	19.48	3.49	22.97
15	9.38	5.95	15.33	4.18	19.52
Med.	9.59	8.80	18.03	3.91	22.12

Table 5

Response Latencies for Lesson 3 of Unit II
on First Presentation

(Average frame of this lesson for each of five latency categories; 7 Ss)

S#	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1	16.73	13.55	30.29	4.96	35.25
2	11.86	7.31	19.17	4.94	24.11
3	17.05	11.25	28.30	4.44	32.75
6	13.56	8.71	22.27	4.24	26.52
7	12.09	9.60	21.68	5.47	27.15
9	14.28	7.63	21.95	3.22	25.20
15	12.70	7.27	19.97	4.75	24.72
Med.	13.56	8.71	21.95	4.75	26.52

Holland-Skinner revised program (10 per cent error), the classroom value was only 2 per cent error. Similarly, while the error rate range for the former was from zero to 3/4, for the latter it was from zero to 14/84. Per cent error and range both differ by a factor of four or five.

Both kinds of error rate data suggest that this particular lesson is somewhat less difficult than the psychology program of Skinner and Holland.

In addition, although the experimental group was small, the discrepancy between error rate in the classroom and that in the relatively more cheat-proof experimental situation may be indicative of a complication in the practice of presenting programmed material in booklet form to classroom students as a means of testing material for purposes of revision. Also, not only may error rate be underestimated, but, in view of the degree of discrepancy, certain kinds of errors may more often be overlooked.

Another way that Lesson 3 of Unit II may be described is in terms of the amount of material on the average frame. The average number of symbols was found to be approximately 128. Each letter, punctuation mark, blank, instance of underlining, special arrow, etc. was considered a symbol.

The correlation between the amount of material of the frames of Lesson 3 of Unit II and the median latency, in the case of each of the five kinds of latency, was considered in terms of Pearson product-moment correlation coefficients, each of the five based on 66 pairs of symbol-median latency. While some of these latency medians were based on all seven Ss, on a few of these 66 frames, medians were based on only 3 Ss. Most medians were based on some number between 3 and 7.

The correlation coefficients for Read, Answer, (Read-plus-Answer), (Score-&-Turn), and Total Frame time were 0.83, 0.03, 0.61, -0.08, and 0.57 respectively.

There are several interesting points concerning this finding. The high correlation between amount of material defined in terms of symbols and reading suggests, again, that Lesson 3 of Unit II was a rather easy lesson (approximately 69 per cent of the variance of these two factors is common variance). In fact, the correlation between reading time and amount of material may be a useful alternate among ways of assessing difficulty in programmed material. Here, the 31 per cent of the variance in reading time not accounted for by amount of material must be accounted for in terms of arrangement of material and various outside influences on S's reading. To the extent that amount of material accounts for reading time, the arrangement of that material is simple.

Incidentally, on the review test of Lesson 3 of Unit II (which was much more difficult in terms of error rate), the correlation, given in terms of a Spearman rank correlation coefficient, was 0.26 between amount of material, Symbols, and (Read-plus-Answer) median latencies. Only about 7 per cent of the variance between amount of material and reading time is common variance; there is room for difficulty or arrangement of material to account for relatively more of reading time.

The number of blanks on a frame of material of Lesson 3 of Unit II ranged from 1 to 9. A complication here and in the case of counting symbols is that when a spelling lesson was superimposed in order to avoid future difficulties with a response term, each space indicating a separate letter was considered a blank. The average number of blanks for a frame of material of 3-II lesson portion was approximately 3 blanks. About half (32 of the 66 frames) of the frames of this lesson required only one word as an answer. Sometimes these required several separate responses, as when missing letters had to be supplied.

Appendix A helps illustrate item content while Appendix B presents the complete review test for Lesson 3 of Unit II.

A comparison of the average latencies of Table 5 with those of Table 2 suggests that differences between Ss in this study are not as great as differences associated with lessons, although this comparison is complicated by the fact that there are 7 Ss and 8 lessons.

There is one way in which Table 5 may be related to the first-version Skinner-Holland psychology program. Skinner and Holland reported an overall median time of 15 minutes on a first cycle, first presentation, of a lesson disk which usually consisted of 29 frames of material (Skinner & Holland, 1958, pp. 42-43). The per cent errors for this more difficult version of their program, it may be recalled, was 20.1. The fifteen minutes, 900 seconds, divided by 29 frames, as the most probable number of frames, gives a total frame time which can be compared with the 26.52 of Table 5, the median based on the averages of 7 Ss. It is approximately 31 seconds.

In comparing programs, it would be interesting to see if the relative proportions of the components of Total Frame were maintained from program-to-program. For example, if the same kind of latencies were obtained in the case of the first-version Skinner-Holland program as were obtained for Lesson 3 of Unit II, would Read, Answer, (Read-plus-Answer), and (Score-&-Turn) be, approximately, (31/26.52) x 13.56, (31/26.52) x 8.71, (31/26.52) x 21.95, and (31/26.52) x 4.75, respectively?

Tables 6 and 7 are based on summarizing medians such as are given in Table 5 (13.56, 8.71, etc.) which appear in row 3-II. These tables are for first and second presentations respectively.

These tables are difficult to interpret since the difference due to lessons is confounded with differences due to Ss. However, one characteristic of 5-II may help account for the fact that, as Table 6 indicates, Answer times were greater than Read on this particular lesson on first presentation. Only on this lesson were Answer times greater. The average frame of Lesson 5 of Unit II required 3.6 response elements, words or letters given in response to blanks or direct questions. (Lesson 3 of Unit II required only 3, actually slightly less than 3.) For 5-II the range of blanks was from 1 to 16, a much greater range than the 1 to 9 of 3-II. In the case of 5-II, subject characteristics as well as lesson characteristics may be responsible for the greater Answer time

Table 6

Typical Response Latencies for Each Lesson for First Presentation

(Median based on average frame of a lesson for all Ss completing that lesson; five categories of latency and 13 lessons; median of these medians given below for each kind of latency, in bottom row)

Lesson	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1-II	13.43	12.58	25.10	6.28	32.30
2-II	15.23	10.58	25.62	4.10	29.48
3-II	13.56	8.71	21.95	4.75	26.52
4-II	13.83	9.24	23.39	3.64	27.12
5-II	9.91	13.10	22.51	4.29	26.42
1-IV	16.37	11.78	30.27	5.42	33.11
2-IV	14.22	8.57	22.34	4.51	26.87
1-VII	18.76	9.09	27.93	7.17	35.19
2-VII	12.42	7.74	20.15	4.90	25.04
3-VII	13.21	8.13	20.66	4.44	24.94
5-VII	16.16	8.83	21.86	5.50	28.67
6-VII	16.30	8.95	22.36	5.13	26.94
7-VII	12.00	10.70	23.02	5.23	27.95
Med.	13.83	9.09	22.51	4.90	27.12

Table 7

Typical Response Latencies for Each Lesson

for Second Presentation

(Median based on average frame of a lesson for all Ss completing that lesson; five categories of latency and 12 lessons; median of these medians given below for each kind of latency, in bottom row)

Lesson	Read	Answer	Read-plus-Answer	Score-&-Turn	Total Frame
1-II	----	----	15.79	----	-----
2-II	11.54	10.65	22.19	4.88	26.12
3-II	9.38	5.95	15.33	4.18	19.52
4-II	9.93	8.85	18.78	3.22	22.00
5-II	8.19	9.15	17.34	2.71	20.06
1-IV	12.78	10.97	23.67	4.56	27.58
2-IV	14.82	9.10	23.84	4.68	28.53
1-VII	9.66	3.46	13.13	2.71	15.84
2-VII	8.65	8.65	16.32	3.65	21.26
5-VII	12.00	7.68	17.01	3.49	23.25
6-VII	14.14	9.03	20.43	4.73	24.90
7-VII	8.99	10.20	19.19	3.67	22.86
Med.	9.93	9.03	18.06	3.67	22.86

relative to Read, however.

The medians based on the averages presented at the end of Table 3 and the similar five medians at the end of Table 6 tell what is characteristic performance on the materials as a whole studied by the group of Ss as a whole. These are two alternate ways (in terms of subjects and in terms of lessons) of defining typical response latencies, and there is rather good agreement between these ways. These comments also apply to Tables 4 and 7 which present the same kinds of information for a second presentation of lessons.

Perhaps one value of these tables is that they permit considering either a lesson (such as 3-II) or a subject (such as Subject #6) in the context of other lessons or Ss. This should help in evaluating the typicality of lesson or S.

Intercorrelations of Read, Answer, and S-&-T

Subject #6 also provided data, from the eight lessons which he completed, that is useful in suggesting relationships between three of the more independent latency measures--independent in the sense of measurement and mathematical operations. The nature of these relationships is suggested by the values of Table 8 which are product-moment correlation coefficients.

Table 8 gives the lesson and the number of frames associated with the three kinds of pairs of scores: Read-Answer, Read-(S-&-T), and Answer-(S-&-T). Only data obtained on a first presentation of these lessons is considered.

Table 8 suggests that, for S #6 at least, there is a slight tendency for longer Read times to be associated with shorter Answer times. Score-&-Turn time seems to increase as Read time increases. Often when there is more material to read, there are more blanks to be filled in; and, thus, more time is required for scoring. Also, difficult material takes longer to read and to score, partly due to the time spent pondering the items missed by S. Answer time seems more closely related to S-&-T. The more time spent in overt responding, the more time spent in scoring and turning to the next frame of material.

Lesson 3 of Unit II is slightly atypical in terms of a slight positive correlation between Read and Answer latencies. For most lessons the correlation between Read and Answer latencies is negative. That this characteristic of 3-II is not due to a peculiarity of S #6 is suggested by Table 9.

The general format of Table 8 applies to Table 9. However, for each kind of pair of latency values, correlation coefficients (product-moment) are given for both the lesson portion and the review portion. The number of pairs of observations for each are given: lesson/review. The number of pairs of observations for all review test data considered was 20. Two Ss do not provide review test data from Lesson 3 of Unit II. Lesson 7 of Unit VII did not have a review test.

Eleven different Ss provided data based on the two lessons, 7 Ss, data based on 3-II and 4, data based on 7-VII. The data of S #6 given for 3-II in Table 8 appears again in

Table 8

Intercorrelations of Read, Answer, and Score-and-Turn Latencies for Subject #6 on First Presentation of Lessons

Lesson	N	Read-Answer	Read-(S-&-T)	Answer-(S-&-T)
1-II	60	-0.06	0.10	0.28
2-II	49	-0.02	-0.10	0.36
3-II	62	0.10	0.13	0.28
4-II	72	-0.16	0.39	0.04
5-II	52	-0.19	0.09	0.05
1-VII	80	0.00	0.03	0.18
2-VII	78[a]	-0.26	0.06	0.45
3-VII	45	-0.07	0.14	0.17

[a] Two observations less for two of these.

Table 9

Intercorrelations of Read, Answer, and Score-&-Turn Latencies for Lesson and End-of-Lesson Review Test on First Presentation for Two Lessons

Lesson 3-II

S#	N	Read-Answer		Read-S-&-T		Answer-S-&-T	
		Lesson	Review	Lesson	Review	Lesson	Review
9	61/20	0.07	-0.31	0.05	-0.22	0.26	0.22
1	56/20	0.12	0.18	0.01	-0.21	0.48	-0.14
15	60/20	0.04	-0.50	-0.86	0.15	0.29	0.33
3	63/-	0.54	----	0.64	----	0.87	----
6	62/20	0.10	-0.24	0.13	0.20	0.28	0.10
7	47/20	-0.13	-0.02	0.62	-0.19	-0.06	-0.20
2	35/-	-0.19	----	0.02	----	-0.06	----

Lesson 7-VII

4	80/-	0.17	----	0.13	----	0.44	----
5	79/-	0.45	----	0.56	----	0.68	----
8	60/-	0.08	----	0.28	----	0.40	----
11	80/-	0.28	----	0.38	----	0.66	----

the row for S #6 in Table 9. The values of his correlation coefficients fall well within the range of values for the various columns. Only in the case of 0.20 for Read with S-&-T on the review test, as Table 9 indicates, is a value of his an extreme value relative to other Ss. This further justifies considering S #6 a typical S.

As far as lessons are concerned, 3-II and 7-VII seem unlike the other lessons presented in Table 8: Read is positively correlated with Answer. Otherwise, they are similar; Read and Answer are both positively correlated with Score-&-Turn. Table 9 also shows that the highest values for the correlation coefficients of Answer with Score-&-Turn are in the case of 7-VII. Lesson 7 of Unit VII was much more difficult than 3-II as it had not been revised on the basis of classroom testing.

First and Second Presentations

Table 10 gives product-moment correlation coefficients for each of the five kinds of latency values on a first and second presentation of a lesson, usually with a week between presentations. These coefficients are given by S and by lesson. While most of these lessons did not have review tests, those that did had the correlation based only on the lesson portion.

By inspection it seems clear that for most kinds of latency, the value on a first presentation is somewhat predictive of that obtained on a second presentation. Whereas in the case of R-plus-A all of the correlation coefficients

Table 10

Correlation Between First and Second Presentations for Lesson Only

(Values Pearson product-moment correlation coefficients; N reduced slightly in some latency categories due to missing data)

S#	Lesson	N	Read	Answer	R-plus-A	S-&-T	Total Frame
3	1-II	66	----	----	0.73	----	----
3	5-II	52	0.52	0.81	0.82	0.84	0.84
4	5-VII	36	----	----	0.82	----	----
5	1-IV	38	0.80	0.80	0.79	0.61	0.84
5	2-IV	50	0.49	0.25	0.75	0.38	0.82
5	5-VII	79	----	----	0.84	----	----
5	6-VII	25	----	----	0.83[a]	----	----[a]
5	7-VII	45	0.52	0.62	0.57	0.06	0.49
6	4-II	72	0.57	0.52	0.81	0.36	0.81
6	1-VII	80	0.45	0.37	0.58	0.29	0.61
7	2-VII	31	0.52	0.73	0.63	0.44	0.57
8	5-VII	80	0.53	0.85	0.65	0.54	0.69
8	6-VII	38	0.23	0.70	0.34	0.59	0.50
10	2-II	52	0.83	0.84	0.89	0.05	0.78
11	1-IV	70	0.14	0.51	0.49	0.34	0.52
11	5-VII	78	----	----	0.66	0.01	0.63
11	6-VII	69	----	----	0.82	0.38	0.82
11	7-VII	80	0.74	0.67	0.80	0.57	0.80
15	3-II	60	0.63	0.59	0.63	0.08	0.65
Median:			0.52	0.67	0.75	0.38	0.69
Range:			.14-.83	.25-.85	.34-.89	.01-.84	.49-.84

[a] Total Frame on First Presentation with Read-plus-Answer on Second Presentation

would be considered statistically significant at the 5 per cent level at least, and are high enough to have practical significance as well, some of the S-&-T values approach zero.

The greater instability of S-&-T values from first to second presentation is reasonable when it is recognized that much of the variance in S-&-T values may be due to errors which may be avoided on a second presentation. Subject #3 made only one error, according to her scoring, on the first presentation of Lesson 5 of Unit II. She did not score any errors on the second presentation. The correlation between her first and second presentation S-&-T values was the highest observed.

Subject #5, by way of contrast, made 12 errors on the first presentation of Lesson 7 of Unit VII, but only 2 on the second. This helps account for the unusually low correlation (0.06).

The lowest value obtained in the case of first and second presentation S-&-T correlation coefficients is that of 0.01 for S #11 for Lesson 5 of Unit VII. On the first presentation, this S scored 14 errors, on the second, 11. However, only 5 of his errors were made on the same frames of material on both presentations.

A similar explanation in terms of decreased common variance due to elimination of errors or their changed location in the material can not be readily extended to other latency categories such as Answer. Not only is there insufficient data relevant to such issues, some of the data obtained is an obstacle to this interpretation. For example, the fact that first and second presentation Answer latencies have a relatively high correlation in the case of Subject #5 for 7-VII, 0.62, while S-&-T is low, 0.06, and that errors decreased from 12 to 2 from first to second presentation should not be overlooked if one is tempted to relate error to degree of first and second presentation Answer latency correlation.

While Total Frame values seem to be relatively stable and yield correlation coefficients of greater magnitude, as a whole, than any other type of measure, Read values are of greater interest as definitions of behavior which may possibly be controlled by way of supplementary reinforcement. The first to second presentation instability of these values was second only to S-&-T. It should be worthwhile to make an effort to understand this better before going on with an attempt to modify reading behavior.

Latency Related to Error

Table 11, by inspection, indicates that larger latency values, in the case of each of the five kinds of latency, tend to be associated with errors to a greater extent than is true of smaller latencies. While several different kinds of sign tests are possible, depending upon one's interests and choice of assumptions, overall tests seem reasonable for each of the five latency categories. All data is considered; each plus or minus is given equal weight. This latter would not be done in the case of making the lesson the unit rather than the frame. Since frames seem as independent as lessons, the more sensitive overall sign test based on frames would seem justified.

Table 11

Summary of Sign Data for Five Categories of Latency on First Presentation

(Individual frames considered; if median of latencies associated with errors greater than median for correct-answer latencies, that frame assigned plus)

	Read		Answer		R-plus-A		S-&-T		Total	Frame
Lesson	plus	minus	plus	minus	plus	minus	plus	minus	plus	minus
1-II	12	9	15	6	17	5	23	0	21	2
2-II	9	5	9	5	14	3	14	0	13	1
3-II	18	7	15	11	20	4	24	1	22	4
4-II	11	1	5	9	10	3	14	0	13	1
5-II	8	6	10	6	8	7	14	1	12	4
1-IV	9	3	7	5	9	4	11	3	8	5
2-IV	4	3	4	4	2	6	7	1	5	3
1-VII	8	6	10	5	11	4	13	1	11	4
2-VII	12	10	14	10	14	10	21	2	16	8
3-VII	6	7	10	4	9	6	14	0	10	4
5-VII	--	-	--	-	26	17	19	10	25	8
6-VII	--	-	--	-	6	1	6	1	7	0
7-VII	13	5	9	8	8	9	15	3	11	7
Total:	110	62	108	73	154	79	195	23	174	51

Accordingly, each test would indicate that the relation between latency and error was significant at, or beyond, the one per cent level, as far as statistical significance is concerned.

All individual frames of material of the 13 lessons studied were considered, and each frame on which at least one S made one or more errors is a basis for the results presented in Table 11.

For each of the five categories of latency, the median was determined for latencies provided by Ss who made an error on a given frame. This median was compared with that based on latencies provided by Ss who did not make an error on that frame. For each frame on which the median associated with errors was greater, a plus was assigned.

A complication in this analysis is that Ss who contributed longer latencies often contributed more errors. It is thus not clear whether or not longer latencies are associated with the behavior of making an error or are associated with a subject characteristic, or both. The results in Table 11 need to be considered in relation to those of Table 12.

Whereas Table 11 presents findings based on holding materials constant, Table 12 displays results when individual differences are held constant. Consider a given S. A lesson completed by S provides an average value in terms of each of the five categories of latency. Each of his errors, provided they are not associated with a latency value exactly equal to the average, may be classified as falling either above or below that average. If more errors fell above his average for

Table 12

Summary of Sign Tests for Five Categories of Latency on First and Second Presentation

(Latencies for frames with errors plus if above, minus if below, S's average of each lesson; when more of S's lessons plus, that S assigned a Plus)

First Presentation

	Read	Answer	(Read-plus-Answer)	(Score-&-Turn)	Total Frame
Plus	10	4	10	16	13
Minus	1	8	2	0	1
p	0.012	(0.39)	0.038	0.00003	0.002

Second Presentation

	Read	Answer	(Read-plus-Answer)	(Score-&-Turn)	Total Frame
Plus	4	2	7	7	7
Minus	1	2	1	0	0
p	0.376	1.00	0.07	0.016	0.016

a lesson for a particular category of latency, a plus was assigned. Since most Ss completed more than one lesson, each S could be assigned several plus or minus signs. When more lessons received a plus than a minus, this was summarized for that S by assigning him a Plus. These capitalized Plus and Minus signs are independent in the sense that there is only one for each S.

While material was held constant in the previous analysis, it is not held constant in this one. If frames of material associated with longer latency also tend to provide more opportunity for error, say, through requiring answers with a greater number of parts, longer latencies may be associated with a characteristic of material rather than with the behavior of making an error. Or, again, the longer latencies may be due to both of two confounded factors. The results of Table 11 complement those of Table 12.

Table 11 indicates that although the relationship between Answer latency and error was found to be statistically significant at approximately the one per cent level (mentioned elsewhere above), this longer-latency-greater-error-rate relationship was less strong than relationships between other kinds of latency and error, such as Total Frame, R-plus-A, etc. When S, rather than material, is held constant, as in Table 12, the relation observed between Answer and error suggest that Answer latencies are not systematically associated with errors.

While scoring an error in terms of operating a scoring mechanism took little time, there was usually some pondering of an incorrect answer in the case of S-scored errors. In both Table 11 and Table 12 the results relating latency to error seem most clear-cut for S-&-T. Total Frame was next as a measure associated with error. Again, this is borne out by both tables. Table 11 suggests that R-plus-A would rank next in terms of degree of association with error and then be followed by Read. There are fewer observations based on Ss, of course, than based on frames involving at least one error. Therefore, Table 11 has more to offer toward establishing such a ranking, although this is not to suggest that there is any real disagreement with such ranking established on the basis of Table 12.

Perhaps the lack of relationship between error and Answer latencies should be considered in somewhat more detail. It may be recalled that Answer deals predominately with overt answering, at least if S followed his instructions. Failure to answer resulted in zero Answer latency. This definition of Answer has not been especially troublesome in specifying, e.g., the average amount of time spent actually writing an answer, from first start to final completion. It stands, however, as a potential complication of analyses such as were reported in Tables 11 and 12. (The fact that few frames were not answered at all and that there were few zeros as a result of faulty answer-switch operation is also relevant, here.)

This problem has been dealt with directly by repeating the analyses reported in Table 11 and 12, the second time with zero values excluded. Table 13 may be related to the plus and minus Answer latency column of Table 11. Such a column is given here as a row; all zero values were excluded in arriving at these results.

Table 13

Summary of Sign Data with Zero-Answer Values Excluded

	Unit II					Unit IV		Unit VII			
Lesson:	1	2	3	4	5	1	2	1	2	3	7
plus:	15	8	15	4	10	7	4	9	9	7	9
minus:	6	4	9	7	6	3	4	4	7	3	8
Total:	plus, 97; minus, 61										

Again, the same conclusion may be drawn, with somewhat greater confidence due to the more refined analysis and due to somewhat greater statistical significance.

When the analysis such as was reported in Table 12 is repeated with zero Answer latencies excluded, there is no greater evidence for an association of latency with error: (1) First presentation plus, 4; minus, 7; (2) Second presentation plus, 1; minus, 2.

Table 14 contains the results of one other analysis concerning the relationship between latency and error. The number of classroom students who made an error on each frame of material presented as an end-of-lesson review test was established, based on classroom use of the material in booklet form. A clerical assistant scored each test according to rather objective standards. (These tests with the answers which served as her key are given in Appendix B.)

Table 14

Rank Correlations of Classroom Errors with Median Latencies and Amount of Material

Lesson:	1-II	2-II	3-II	5-II
Classroom Errors with R-plus-A Latencies:	0.83	0.58	0.58	0.72
Classroom Errors with Number of Symbols[a]:	0.24	0.22	0.09	0.09
Number of Symbols[a] with R-plus-A Latencies:	0.21	0.21	0.26	0.22
Number of Classrooms as Sources of Error Data:	5	5	4	3
Total Number of Students from These Classrooms:	129	127	102	62
Range of Values of N for R-plus-A Medians:	3-5	2-3	5	3-4

Note.- For N equals 20, r_s equals 0.377 at the 5 per cent level and r_s equals 0.534 at the 1 per cent level (one-tailed test), Siegel (1956) p. 284.

[a] Amount of material (or Symbols) was defined by the total number of letters, blanks, marks of punctuation, special lines, arrows, etc. per frame of material.

Three to five Ss provided Read-plus-Answer latencies in the case of each review test, although some of the Ss provided data on more than one test. The median of these R-plus-A values for the three to five values for each frame was obtained. Consequently, each frame of a review test had a pair of values associated with it: Classroom errors and median R-plus-A latency. A Spearman rank correlation coefficient was based on the ranks of such pairs of values. Each review test had the same number of frames; the four correlations were each based on 20 pairs.

Again, if a relationship were found between the latencies of Ss and the errors of classroom students, it might be largely accounted for in terms of characteristics of materials. Frames that present much material may both take longer to complete and lead to more errors. Amount of material (defined in terms of a simple sum of the number of letters, blanks, marks of punctuation, special lines, arrows, etc. per frame, as before in the case of 3-II lesson portion) was considered in a correlation--this time Spearman rank--with classroom errors. Amount or symbols entered the analysis just as median latencies had, as a value paired with classroom errors. Amount was next related to median R-plus-A latencies. These results are all presented in Table 14.

Table 14 indicates that the R-plus-A medians, based on as few as two or three Ss in the experimental group, were significantly associated with the errors of another, larger group of students who completed the same material in booklet form.

These students came from several different classrooms, from three to five different ones. The number of students providing error data varies for the four correlations from 129 to 62. As far as the experimental group was concerned, the number of cases providing a latency value from one frame to the next varied. Therefore, the values given in the lower row of Table 14 indicate a range, rather than a single value, for the number of Ss.

Table 14 suggests that the prediction of error rate may benefit from taking into account R-plus-A latency values. In the case of each of the four correlations, more than a third of the variance between latency and error is common variance.

Table 15 gives the values of Spearman rank correlation coefficients in the case of individual Ss. This table indicates that in some cases R-plus-A values of an individual S have nearly 50 per cent of their variance common to classroom error variance, that the R-plus-A latencies of some individual Ss could be as predictive of error rate as group medians. However, Table 15 also indicates that the degree of association between R-plus-A latency and error rate, for the individual S, varied from lesson-to-lesson. This may have been due to lesson characteristics, but it may be a result of changes of subject behavior from one occasion to the next, or both.

The prediction of error rate by means of R-plus-A latencies does not seem greatly dependent upon amount of material. That is, while there is a slight tendency for error rate to increase with an increase in amount of material and for

Table 15

Rank Correlations of R-plus-A Latencies of Individual Subjects with Classroom Errors on Review Test

Subject #	Lesson	Frames	Correlation Coefficient
1	1-II	20	0.78
1	3-II	20	0.30
3	1-II	19	0.65
3	2-II	16	0.70
3	5-II	20	0.78
6	3-II	20	0.49
6	5-II	19	0.65
7	3-II	20	0.67
7	5-II	17	0.48
9	3-II	20	0.30
10	2-II	20	0.69
13	1-II	19	0.52
15	2-II	19	0.45
15	3-II	20	0.63
16	5-II	17	0.51

R-plus-A latency to increase with amount, this tendency is not as strong as the association between R-plus-A and error rate. As a common third factor, amount of material does not completely account for the observed relationship. In no instance does amount of material yield a Spearman rank correlation coefficient of greater than 0.26 in the case of correlation with either classroom errors or medians of R-plus-A latencies.

This finding based on the more difficult material of the review tests constrasts with the previous finding in the case of the lesson portion of Lesson 3 of Unit II. It suggests that it is likely to be easier to predict item difficulty by means of latency where there is relatively greater difficulty. As programs are revised and made easier, relatively more of the latency value comes to be due to amount of material in all measures involving reading, thus all latency measures. For lesson material of 3-II, it may be recalled, a product-moment correlation coefficient of 0.83 was obtained which was based on 66 pairs of median Read latencies with amount of material.

Amount of material will need to be taken into account in any attempt to revise programs on the basis of latency data. Furthermore, this may be complicated by the fact that, as amount increases, often there will not be a proportional increase in reading time. As more material is added to a frame, the material added last often may be proportionately easier as there is more context aiding its interpretation.

Whether or not error rate was more closely associated with amount of lesson or review material could not be considered.

Error data for the kinds of analyses carried out with review tests were not available in the case of the easier lesson material.

Until some of these issues are investigated directly, latency data would seem most useful in assessing difficulty of relatively difficult material, such as was found on the review tests. However, as the difficult portions of review test material come to be recognized, the programmer has some suggestion, on the basis of the content of the difficult portions, concerning frames that need to be changed in the preceding lesson. What he would lack would be knowledge concerning whether these lesson frames were defective on the basis of content, whether they caused the student little difficulty but were still not helpful in teaching what the programmer intended, or whether they failed to teach adequate content which simply was not yet put into the form of sufficiently prompted items. Recognizing inadequate frames, the programmer still would not know what was wrong with them.

Some Applications to Program Analysis

Some of the latency data was applied to program analysis in an attempt to discover ways in which this data was immediately useful. This was certainly not an exhaustive or systematic study. It is offered in the sense of some suggestive observations.

It was mentioned earlier that latency data may complement error rate data. An example of this involves a frame of material which required an unusual amount of time for a number of students. The content of one error may explain to the programmer just what accounts for the confusion which rarely resulted in error but usually resulted in an unusual amount of time being spent. A frame of material which illustrated this was one that was intended to prompt by indicating with spaces the number of letters of the correct response. Two spaces were typed too close to one another. When a student gave an answer which was spelled incorrectly and contained one letter less than the correct answer, it led to the discovery of this flaw. One error in spelling would not have been given much consideration if the latency data had not already aroused some interest.

Sometimes a student may miss an item and the programmer may wonder: "Did this student really understand topic X or was it topic Y that troubled him?" By considering the latency record in relation to the program content, the programmer may discover that for items which deal with topic X, latencies on successive frames tend to become reduced. On frames that deal with topic Y, greater latencies in the case of one or two frames late in the development of this topic may suggest confusion just before this topic was concluded. A prompt that favored recall of topic Y could reduce difficulty at this point, if the difficulty of this student was a typical difficulty. The important content of the better-mastered topic X would be recalled without a prompt. The prompt, which repeated important content of topic Y, could later be vanished. A more economical revision might be to return to topic Y and improve the frames in the latter part of the development of this topic.

Often errors are made on simple frames of material when a topic is approached from some novel frame of reference. The student who has been making few errors may simply read the frame carelessly. If the Read latency value was high on this frame, such an interpretation would not seem justified. The student who is asked to give a semicolon and who gives a colon instead, spending twenty seconds on this frame, does not give the same impression of carelessness as the student who spent five seconds making that same mistake.

In applying latency data, there may be a temptation to reduce the data of all students to an average or median value and then consider material in terms of associated latencies which deviate from such norms. While this may be a reasonable procedure, one should not overlook the value of relating records of selected individual cases to the analysis of a program. Both median values and values from selected individual records may prove useful in the analysis of a program. In the present study, Subject #12, the foreign student who had unusual difficulty with English, provided some of the most useful data as far as detecting some of the characteristics of material that often were not necessary and that often were wasteful of time for most students. For many students this waste was slight and not easily recognized.

Sometimes the use of hints introduced extra reading time only to fail their intended function as a result of a vocabulary problem or a general ambiguity. Statements which were used as stated hints often were no more helpful as prompts than other kinds of prompts. Usually such hints required more time.

Students seemed to answer more quickly the single blank than the blank broken up into letters. The use of spaces for letters making up the answer became wasteful of student time for all students who could give the correct answer without such an aid. Even students who were relatively confident of their answers probably found it difficult to resist taking the additional time to check and just make sure.[2]

In studying the records of S #6 and S #7 in relation to the content of Lessons 2 and 3 of Unit VII, there seemed to be some qualitative differences in performance. Subject #6, who was much faster, tended to make mistakes in terms of faulty use of concepts taught in the program. His performance often indicated the kind of mistake one would make if he did what he had been encouraged to do by the program earlier. It rather often pointed up faults of the program in terms of subject matter in contrast to the performance of S #7.

Subject #7 seemed to make errors due to careless or faulty reading. Also, he might give an answer inappropriately that had been repeated several times on previous frames. Remote program content did not seem to be as helpful in accounting for his, more frequent, atypical responses. Also, instead of faulty applications of relevant concepts, his errors seemed due to a more heterogeneous set of influences.

---

[2] One fourth of the frames of 5-II lesson material involved the use of spaces for letters. However, 5-II was similar to 3-II in this respect. Therefore, this characteristic of the material cannot completely account for the longer Answer times than Read times of 5-II.

Covert- Compared with Overt-Answering

During the course of the study an opportunity arose for obtaining some data related to the issue of evaluating the probability of covert responses on the basis of latency data. A subject who was understood to be working on a lesson, reading and answering covertly, provided a record which was related to records provided by other Ss who had given overt answers for that same lesson.

Subject #7 completed 4-II using the modified machine after having been exposed to the same material in the classroom in booklet form. S believed he could benefit from more study of this lesson, but he was not sure the gain would fully justify the time. Consequently, he chose to work the lesson without writing his answers.

His latency values were considered in terms of Pearson product-moment correlation problems which related his Read-plus-Covert-Answer latency values to the median latencies based on the three Ss who had completed the same material under the usual conditions of giving overt, or written, answers.

It would seem probable that Subject #7 actually did answer the material covertly if his pattern of latencies on the frames of 4-II closely resembled that of the Ss who answered overtly. The purpose of considering these problems simply is to demonstrate one method for assessing the probability of covert responding when such responding may or may not have occurred. Correlation coefficients which can be related to some standard of chance expectancy make possible one assessment of degree of agreement between a pattern of latencies, possibly based on covert responding, and one definitely based on overt responding.

Each of the 72 frames of 4-II, which was the only lesson of Unit II which did not have a review test, had two values associated with it in successive correlation problems. First, consider Subject #7's Read-plus-Covert-Answering values as making up one member of each of the 72 pairs of values. The median based on the latencies of the three Ss who wrote their answers for the same frames was then the other member of the pair of values.

Three correlation coefficients were computed: (1) Median Read with Read-plus-Covert-Answer, (2) Median Answer with Read-plus-Covert-Answer, and (3) Median Read-plus-Answer with Read-plus-Covert-Answer. Respectively, r equaled 0.58, 0.57, and 0.76. Rarely would such a set of values be obtained by chance. In terms of a one tailed test, a value as large as 0.38 for 70 degrees of freedom would be expected about 5 times in 10,000 on the average. Furthermore, the latency suspected of containing the covert answering component shows the highest correlation with the median latency which does not contain merely Answer or Read--the highest correlation is between overt R-plus-A and covert R-plus-A. (Incidentally, a correlation coefficient of 0.46 was obtained in the case of Subject

#7's Score-&-Turn values correlated with the group median Score-&-Turn values. Even the scoring of covert answers provided a latency pattern which afforded better than chance agreement with that based on overt responding! This is undoubtedly due largely to the reading component, although pondering may be similar too.)

For some future, refined study of overt compared with covert answering as a programming issue, it may become desirable to make certain that covert responding actually is occurring in one of the experimental groups. Here is one method for establishing the presence or absence of covert responding that permits some consideration of the role of chance in making the decision.

It may be of interest to note, since Subject #7 was singled out for comparison with Subject #6 previously in connection with possible qualitative differences in program-working behavior, that the high agreement of latency pattern of this S with that of group medians does not suggest any qualitative difference in latency pattern. Perhaps individual differences may be expected to show up more in relation to program content than in characteristics of latency--at least until latency analysis is further refined.

Supplementary Reinforcement

A preliminary attempt was made to apply bonus point reinforcement to Read latencies which fell within a specified range of values. On the basis of a first presentation of a lesson, ranges of Read values were established that applied to each frame of the lesson. Latencies for reading which were approximately equal to or shorter than those for the same frames on a first presentation would occasion the flashing light. A bonus point would be delivered as an increment on the counter in view of S, at the teaching machine, when frames were completed on the second presentation for which Read latencies fell within "improvement" intervals--unless an error had been scored on such frames. Responses of reading too quickly, indicated by an implausibly quick overt answer, should be excluded by the specific definitions of improvement interval.

However, in this preliminary work, the timing circuit which was used to define improvement interval, and to change this definition appropriately from frame-to-frame, may have given rather gross definitions in some instances. Furthermore, these intervals were not monitored by any recording system nor was the delivery of a bonus point related by some record to the frame of the program which immediately preceded its delivery.

The responses of Ss to bonus points can only be given as casual observations: (1) Two of the three Ss reported that they did not find the flashing lights difficult to get used to while the other S (who made more errors than the other two Ss) considered them somewhat more distracting. (2) All three Ss liked the added bonus-point feature, however. (3) Two Ss seemed to respond in opposite manner. One S worked unusually fast; the other S, unusually slowly. (4) The latter S seemed to develop some superstitious behavior ("When lots of items are alike I could get an extra point.") (5) Ss may require special training and staging of such complex schedules. The S who slowed down seemed to have received an early reinforcement according to the adjusting component of the bonus-point schedule. While the long range performance may assume the characteristics expected of a schedule, the time required for stabilized performance may be too great for it to be practical simply to wait for the schedule to "take effect."

Miscellaneous Comparisons

Results are presented in a form which readily permits comparisons implicit in the arrangement of results. For example, latencies will be seen to be reduced in value in virtually all instances when first and second presentations are compared. Irregularities here may be of interest in relation to some special categories. Other materials in Appendix C make possible a comparison of latencies and errors for lessons with those for review tests.

Subject #12, who had special difficulty understanding English, had an average Total Frame latency of 38.15 on the first presentation of Lesson 1-II (as indicated in Table 16 of Appendix C). Three of these frames (data is not given in this much detail anywhere in this report) with greatest latency, 132, 107, and 99 seconds, were also frames on which errors were made. That frames which are wasteful of student time do generally tend to reveal their faults through error is certainly in line with the other results of this study.

Although the data of only one S, that of S #12, should not be considered as conclusive, here is a suggestion that latency data may not be uniquely helpful in discovering frames that present special difficulty for the foreign student. Such frames should be recognized also in an analysis of a program in terms of error rate.

The repetition of lessons seemed to provide possibly interesting results. On what seemed to be unusually substantial frames of material, (frames providing crucial definitions, difficult discriminations, statements and applications of important principles), it was noted that S often spent longer on a second than a first presentation.

While not all frames which are associated with extra time in this manner may be exceptionally useful frames, the more exceptional frames may be more likely to fall within this category. Care should be taken in an application of bonus-point reinforcement that S does not receive reinforcement for slighting important material.

Chapter VI

Discussion and Summary

The correctness of the casual view that longer latencies tend to indicate difficulty was borne out by three lines of evidence. First, when material was held constant, errors tended to be associated with longer latencies for each kind of latency studied. Second, in the case of the individual S, except for overt Answer, errors tended to be associated with his longer latencies. Third, latencies in the case of review test material tended to be associated with classroom errors in a manner not wholly accounted for in terms of a common third factor, amount of material.

One complication, however, was indicated in the case of the easier, lesson material. As material became less difficult, more of reading time could be accounted for in terms of amount of material. Since reading is a component of all latency measures, prediction of error rate from latency data, for all categories of latency, should be improved by taking amount of material into account.

While the relationship between reading time and amount of material could be considered in the case of a lesson, amount of material could not be considered in relation to error rate. There were not sufficient student-scored errors in the case of this easier material to establish a relationship between them and amount of material.

Perhaps future study on a larger scale will help clarify this matter and offer a practical benefit by pointing up a technique whereby the part of latency common to material is separated from that which is, more uniquely, common to difficulty. Basic to such study is error data obtained from many Ss and obtained under relatively cheat-proof conditions. The complication involving amount of material is likely to be only a minor and temporary obstacle to the use of latency data in program analysis.

In view of the close relation between error and latency, as indicated by the analyses of this study, some of the advantages of latency for program revision deserve careful consideration.

First, the preparation of material is costly. While a large number of Ss may be required to establish differences between frames of material in terms of error rate, each S provides some information about each frame in terms of latency data. Early program revision could be accomplished with only a few Ss before a more costly, large-scale test was attempted. As long as subject-matter content was not sacrificed, revision should be profitable which reduced the time spent on a frame of a program. Even if the revising which reduced this time did not also reduce error rate, revision would not be useless. The results of this study suggest, however, that latency data do guide the programmer toward both kinds of benefit.

Second, the greater sensitivity of latency measures recommends them for study of many issues concerning characteristics of programmed materials: techniques of prompting, organization of program content, visual displays of material, and form of prompt or answer.

Third, as analytic categories, latency measures may help in the discovery of qualitative differences in the responses of different individuals to the program. For example, students of the long-reading-time category may be found to respond to certain kinds of prompts involving relatively immediate material. Students with shorter reading times may be found to be responsive to subject-matter covered several lessons earlier.

Finally, latency data may help indicate conditions under which an error was made, e.g., whether or not it is likely to have resulted from careless reading. These data may indicate whether or not a student quickly gave up trying to answer a particularly difficult frame. With further study, a greater range of interpretations may become possible. Latency data may suggest something about such matters as changes in student interest (attentiveness).

Since programs involve many frames of material and require much student time, future investigations of differences in time and in learning in the case of covert- compared with overt-answers may be expected. A way of considering the possibility of covert responding was described. A high correlation between what might be covert responding of a given S and group data in the case of overt responding removes some of the doubt concerning the extent to which that S was responding covertly.

A more general application suggested to E has been that latency might be useful in describing performance so that students who needed it could receive extra help in making the best use of automated materials. Students who might otherwise spend the entire semester making improper use of the teaching machine and the program could be recognized and helped in time to obtain full benefit from the automated instructional situation.

Early in the semester, detail could be obtained routinely indicating which students needed special help. The performance of a student at a modified machine for only a short time could provide a record that an instructor could use diagnostically.

One student might provide a record indicating many short latencies and also many errors. The instructor could show this student how to make better use of the program by slowing down and giving the material more thoughtful consideration.

Another student might provide a record indicating many long latencies and no errors. This student could be encouraged to work faster. He could be instructed not to slow down unless he found that he was making two or three errors per lesson (or whatever seemed a reasonable balance between avoiding a waste of time and of making errors). The student who was working through the lesson too slowly and carefully could be noticed and helped in making more efficient use of the program.

The student who made no errors and who provided characteristically short latencies could either be encouraged to skip ahead or to use a program with somewhat different content.

The student who provided both long latencies and many errors could be given programmed material with many short steps if this was available. He could be encouraged to repeat lessons, although alternate forms of a program which covered the same material would be preferable to repetition of exactly the same programmed material.

Now that relatively inexpensive teaching machines are coming to be commercially available, using microfilm and having push-button rather than mechanical controls, a new avenue is opened to this study. As the student's performance is considered, the programmer could monitor S's work on a lesson rather closely. By starting with the same film in two teaching machines, the one used by the student and the one observed remotely by the investigator or instructor, latency values could be related to frames of the program as they were immediately given. The investigator could observe, in another room, exactly the same material as S, at the same time and for the same length of time. He could observe how S was operating the teaching device. On the basis of such detailed observation of the performance of a student who was known to be making some kind of atypical use of the program or device, suggestions could be made. If the investigator became interested in some peculiar characteristic of S's performance, he could even interrupt S for a moment and question him about why he was working a certain way. The questioning of this kind that was not hampered by a long delay between S's work and the questions concerning it, could also help in the discovery of special subject-matter difficulties.

If the investigator had some special interest in the accuracy of student self-scoring, it could be an inconvenience to have to score the entire lesson. One finding of the present study suggests that errors in scoring would be most likely to be found on frames that were associated with longer than average Score-&-Turn latencies.

The technique of allowing the student to control the frame-by-frame advance of both the program in the student's teaching machine and a duplicate copy in the investigator's somewhat remotely located teaching machine permits several kinds of study of student behavior in relation to program content. In addition, it could help in the revision of the program. At the time a student gave an extremely quick response to a frame, a response that was also a correct response, the investigator might then question him to find out what principle or concept he had related to that frame of material which made it so easy for him to answer. Sometimes it might be discovered that correct answers were given quickly and with considerable confidence for the wrong reasons.

Some frames of material may be characteristically associated with short latencies for students in general. This would suggest that answers to these frames required very little "thinking." If students in general were responding to these frames correctly for the desired reasons, these frames might well be dropped from the program for the sake of economy. Trivial items are, perhaps, less likely than excessively difficult items. Nevertheless, they are certainly a possibility.

The present study, although using relatively unrefined measures, provides a view of greater order than might have been suspected. There is much to recommend further study and the introduction of refinements. The pattern of latencies over frames of material on a first presentation typically bore a striking resemblance to that on a second presentation.

The repetition of lessons has been available as a means of discovering what frames of material are especially difficult by virtue of the repetition of error. As the potential value of more detail is seen, research may more often entail this technique. One incidental observation of the present study was that longer latencies on the second presentation of a lesson often seemed to be associated with important frames of that lesson. Here, review tests may be what is most important in determining performance on the second presentation. If so, performance on the repetition of a lesson may provide feedback for the revision of end-of-lesson review tests.

Branching to different program content, or the skipping of items in a program, could be based on both student-scored errors and latency. Short latencies and few or no errors could be the condition which automatically changed the kind of content or the number of frames of the program. The programmed materials could be kept most appropriate for S at any stage of his development; this could be done automatically if it were worth the cost. The cue-clarity considered in Chapter II might well be brought under the control of student latency.

The suggestion of possible developments deriving from further study of student response to programmed material offers much gain for keeping the final authority in mind. As his role is enlarged, machine design should incorporate what he tells us about the conditions most favorable to programmed learning.

The study of learning needs the techniques and controls of the laboratory, but it also needs the phenomena of the classroom. The teaching machine that is modified for research purposes offers many opportunities which are yet to be explored.

References

Carr, W. J. Self-instructional devices: a review of current concepts. Wright-Patterson Air Force Base, Ohio: Aero-Space Medical Laboratory, Wright Air Development Center, August 1959. WADC Technical Report 59-503. In A. A. Lumsdaine & Glaser, R. (Eds.), Teaching machines and programmed learning: a source book. Washington, D. C.: National Education Association, 1960. Pp. 540-562.

Evans, J. L., Glaser, R., & Homme, L. E. A preliminary investigation of variation in the properties of verbal learning sequences of the "teaching machine type." In A. A. Lumsdaine & Glaser, R. (Eds.), Teaching machines and programmed learning: a source book. Washington, D. C.: National Education Association, 1960. Pp. 446-451.

Ferster, C. B., & Sapon, S. M. An application of recent developments in psychology to the teaching of German. Harvard Educational Review 28: 58-69; Winter 1958. In A. A. Lumsdaine & Glaser, R. (Eds.), Teaching machines and programmed learning: a source book. Washington, D. C.: National Education Association, 1960. Pp. 173-185.

Ferster, C. B. & Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Fry, E. B., Bryan, G. L., & Rigney, J. W. Teaching machines: an annotated bibliography. Audio-Visual Communication Review 8: 1-80; Supplement 1, 1960.

Holland, J. G. Teaching machines: an application of principles from the laboratory. *J. exp. anal. Behav.*, 1960, 4, 275-287. In A. A. Lumsdaine & Glaser, R. (Eds.), *Teaching machines and programmed learning: a source book.* Washington, D. C.: National Education Association, 1960. Pp. 215-228. Lumsdaine and Glaser mention that this was in press in *Proceedings of the Educational Testing Service Invitational Conference*, October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960. The *J. exp. anal. Behav.* source was consulted.

Keislar, E. R. The development of understanding in arithmetic by a teaching machine. *J. educ. Psychol.*, 50: 247-253; 1959. In A. A. Lumsdaine & Glaser, R. (Eds.), *Teaching machines and programmed learning: a source book.* Washington, D. C.: National Education Association, 1960. Pp. 425-436.

Lumsdaine, A. A. & Glaser, R. *Teaching machines and programmed learning: a source book.* Washington, D. C.: National Education Association, 1960.

Meyer, Susan R. Report on the initial test of a junior high-school vocabulary program. In A. A. Lumsdaine & Glaser, R. (Eds.), *Teaching machines and programmed learning: a source book.* Washington, D. C.: National Education Association, 1960. Pp. 229-246.

Siegel, S. *Nonparametric statistics: for the behavioral sciences.* New York: McGraw-Hill, 1956.

Skinner, B. F. & Holland, J. G. The use of teaching machines in college instruction. Final Report to The Fund for the Advancement of Education, August 15, 1958. Part I of this four-part report was the basis for an article by Dr. Skinner in *Science*, October 1958. The four parts appear in Lumsdaine & Glaser, 1960, Pp. 137-172. Although the original was consulted, the practice of citing Part I as Skinner (1958) and the other three parts as Skinner & Holland (1958) was adopted from Lumsdaine & Glaser (1960).

Zimmerman, D. W. Durable secondary reinforcement: method and theory. *Psychol. Rev.*, 1957, 64, 373-383.

PRECEDING PAGE MISSING

Appendix A

Sample frames of material are given which are typical, in many respects, of most of the material of Unit II. These frames involve various response characteristics: from one to four blanks, the selection of one of two correct phrases, missing letters so that a spelling difficulty is avoided by considering troublesome letter combinations early, and a simple yes or no question.

These sample frames are given in terms of error rate: low, moderate, and high. These refer to classroom data which were not obtained under cheat-proof conditions. (In fact, the error rate was about 5 times as high for the experimental group as the classroom in the case of Lesson 3 of Unit II.) Nevertheless, the same frames of material seem to be designated low, moderate, or high in terms of either kind of error rate.

The first three lessons of Unit II, from which these samples were obtained, were responded to by more Ss than other materials. Data were obtained from seven Ss on each of these lessons, although observations were missing from some of these records. Also, some of the same Ss provide data in the case of each lesson.

Five kinds of latency are considered in relation to each frame. The median latency of those associated with correct answers are given relative to the median value of latency associated with incorrect answers (correct/incorrect). In the case of Low Error Rate frames, the median is usually the value of a single latency value in the case of each of the five kinds of latency. Only one error was made on such frames.

There are several reasons for the medians' for Read plus those for Answer not adding to give Read-plus-Answer medians. Sometimes the number of observations was greater in the case of either Read or Answer than Read-plus-Answer since the loss of either kind would also reduce the number for Read-plus-Answer. However, there may have been a greater number of Read-plus-Answer latencies on which to base the two kinds of medians than for either Read or Answer. Early records, or records obtained during faulty action of the answer-switch, sometimes provided an excellent Read-plus-Answer record while failing to provide these latencies separately. Defective switch action meant zero values in the second channel print-out record usually. Read values, or first channel values, were recognized easily as actually Read-plus-Answer values.

Appendix B

All of the review test materials used in this study are given in Appendix B. Here the number of classroom students who made one or more errors on a frame is given as the numerator of the fraction which appears at the right below the frame of material. The denominator is the number of students taking the test in booklet form. The identifying number of Ss taking each review test, their ranges and average values for each of the five kinds of latency, and the number of

E-scored errors with associated latencies which were above or below each of these five averages is given in the third, and last table of Appendix C.

Appendix A

Frames Selected from Lessons of Unit II as Examples of Three Levels of Difficulty According to Classroom Error Rate

(Median values of latencies associated with correct answers compared with median values for incorrect answers for each sample frame of material for five latency categories)

Lesson 1

Low Error Rate

A word that names a ________, ________, or ________ is called a ________. (ans. person, place, thing (any order), noun)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	8/14	16.5/22	24/36	6/22	31/58

Moderate Error Rate

The framework of the simplest English sentence will consist of a ________ and a ________ (part of speech). (ans. subject, verb)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	8/7	26/21	27/17	5/11	43/31

High Error Rate

The framework of a simple sentence may consist of a ________, a ________, and an ________. (ans. subject, verb, object)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	12/9	20.5/23	26/43	6.5/9	39/52

Lesson 2

Low Error Rate

The object of a verb will be either a ________ or ________ (parts of speech). (ans. noun, pronoun)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	4/1	8/28	15.5/29	3/6	21/35

Moderate Error Rate

A direct object ________ the action of the verb. (ans. receives)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	6/5.5	4.5/11	11/16.5	3.5/7	13/23.5

High Error Rate

If an object contains the word to or for before it, the object is (an indirect object, part of a phrase?). (ans. part of a phrase)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	18/51	8.5/2	25/53	2/9	30.5/62

Lesson 3

Low Error Rate

A direct object answers the question what or whom r _ _ e _ _ e d the action. (ans. received)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	5/18	7/5	16/18	3/11	19/29

Moderate Error Rate

He became a captain. He became shrewd.

Is there an object complement or direct object in either of the sentences above? (ans. no)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	5/10	2/3	10/12.5	3/8.5	13/21

High Error Rate

The verbs tastes, smells, feels, looks, and seems may be used as ________ or ________ verbs. (ans. linking, copulative)

Median Latency for Correct/Incorrect	Read	Answer	R-plus-A	S-&-T	Total Frame
	10/15	16/8	21.5/26.5	9/8	35/30

Appendix B

The Four End-of-Lesson Review Tests of Unit II

The following instructions were given at the beginning of each test:

In most of the reviews in this program, you are expected to fill in the blanks. An answer will be counted entirely wrong if it contains an error of any kind - in spelling, omission of an article, or capitalization. You are not to use abbreviations.

If you are asked to write sentences, the sentence will also be counted entirely wrong if it contains an error of any kind - in spelling, capitalization, grammar, or meaning. Hence a sentence will be either correct or incorrect.

All answers have been typed in the blanks; the need for additional alternatives will occasionally be obvious. This was a first testing of these materials.

Review Test for Lesson 1

1. The key words of a sentence are the subject and a verb 29/129
2. The word that answers who? or what? about the verb will be either a noun or pronoun 41/129
3. Words that modify nouns or pronouns are called adjectives 14/129
4. Words that answer the question how , when , or where (why) about the verb are called adverbs . 74/129
5. The object of a preposition is usually a noun or pronoun 12/129
6. A verb expresses being , action , or state of being 15/129
7. Nouns or pronouns may be used as the objects of verbs or prepositions 57/129
8. Prepositions and conjunctions connect elements. 63/129
9. Adjectives modify nouns or pronouns ; adverbs modify verbs , adjectives , or other adverbs . 24/129

10. The basic framework of an English sentence may consist of a subject and a verb or of a subject , a verb , and an object 76/129

11. A complement is a word or group of words which completes the thought of a sentence. . . . 52/129

12. A direct object answers the question what or whom received the action. 81/129

13. A direct object is a complement which completes the meaning and receives the action or is the result of the action. 95/129

14. A noun or pronoun used as the object of a verb or preposition will be in the objective case. 65/129

15. In the pattern subject-verb-object, the verb expresses action 2/129

16. The next frames will be a review of the two sentence patterns. Write a sentence to fit each of the patterns.

subject (noun) verb 8/129

17. subject (pronoun) verb 15/129

18. subject (noun) verb object (noun) . 20/129

19. subject (pronoun) verb object (noun) . 28/129

20. subject (noun) verb object (pronoun) 54/129

Median: 35/129

Range: . . . . 2/129 - 95/129

Review Test for Lesson 2

1. A noun is the name of a person , place , or thing. 7/127

2. A pronoun is a word which stands for a noun . 7/127

3. A verb expresses action , being , or state of being 16/127

4. Adjectives and adverbs modify words. . . . 11/127

5. Prepositions and conjunctions connect elements. 39/127

6. The subject or the object of a verb will be either a noun or pronoun 39/127

7. A verb fits the pattern act/ acts , acted , have acted 89/127

8. To represent past tense the only word in the sentence you change is the verb 4/127

9. A preposition is followed by a noun or pronoun called its object 22/127

10. A simple sentence may consist of a subject and a verb or a subject , a verb of action , and an object 86/127

11. A noun or pronoun used as the object of a verb or preposition is in the objective case. 50/127

12. A direct object follows the verb. . . . 15/127

13. Indirect objects precede the direct objects of verbs. 19/127

14. The subject answers the question who? or what? about the verb. 71/127

15. The direct object answers the question whom? or what? received the action. 49/127

16. Indirect objects can be changed to prepositional phrases by inserting the word to or for before the object. 20/127

17. Using the indicated verb, write sentences to fit this pattern: subject-verb-indirect object-object

Example: to give The city gave the winners a bonus.

to pay 42/127

18. to send 37/127

19. to write 38/127

20. to make 46/127

Median: 37.5/127

Range: . . . . 4/127 - 89/127

Review Test for Lesson 3

1. The use of a word determines its part of speech 15/102

2. A complement is a word or group of words which complete the thought of a sentence 48/102
(alters)

3. A word that limits, qualifies, or describes another element is called a modifier 70/102

4. Any word, phrase, or clause that modifies a noun or pronoun is an adjective in function. . . 7/102

5. Any word, phrase, or clause that modifies a verb , an adjective , or another adverb is an adverb in function. 30/102

6. The position of an adjective may be before or after
(three words)
the word it modifies 25/102

7. The noun or pronoun following the verb in the subject - verb - object pattern will be in the objective case. 20/102

8. The object of a verb or preposition will be classified as a noun or pronoun 56/102

9. A subject complement may be a noun , a pronoun , or an adjective 16/102

10. The simplest sentence pattern consists of a subject and a verb 15/102

11. A simple sentence may consist of a subject , a verb of action , and an object 46/102
(three words)

12. A basic framework of a simple sentence may be a
(copulative)
subject , a linking verb and a subject
(two words)
complement 66/102

13. The verb in the pattern S-V-O expresses action ; the verb in the pattern S-V-SC represents being or state of being 33/102
(three words)

14. A word which follows a linking verb and is equal to the subject is a subject complement . . 62/102

15. A word which follows an action verb and receives the action is a direct object 61/102

16. Write sentences to fit each of the indicated patterns:

Example: subject pronoun - linking verb -
He is

modifier - adjective
very handsome.

subject (noun) - verb - object (noun) 39/102

17. subject (noun) - linking verb - complement (noun) 37/102

18. subject (noun) - linking verb - complement (adjective) . 23/102

19. Use the indicated verb in a sentence which fits the pattern:
subject (pronoun) - (to study) - object (noun) . 40/102

20. subject - (to feel) - complement (adjective) . . 45/102

Median: 38/102

Range: . . . . 7/102 - 70/102

Review Test for Lesson 5

1. The subject of a sentence answers the question who? or what? about the verb 20/62

2. The direct object answers the question what? or whom? received the action of the verb. . 29/62

3. A subject complement is equal to or describes the subject 20/62

4. A subject complement may be a noun , a pronoun , or an adjective 10/62

5. A direct object may be either a noun or a pronoun 3/62

6. Three basic patterns of a simple sentence consist of a subject and a verb ; a subject , a verb , and an object ; or a subject, a verb , and a subject complement 33/62

7. Indirect objects can be changed to prepositional phrases by inserting the word to or for before the object. 30/62

8. A co-ordinating conjunction joins sentence units of equal rank. 12/62

9. Only verbs and verbs , adverbs and adverbs , adjectives and adjectives may be used together in the same compounds. 8/62

10. A simple sentence may contain a compound subject , predicate , object , or subject complement (Elements of a sentence) 41/62

11. A noun or pronoun used as the object of a verb or preposition is in the objective case. 31/62

12. In analyzing an English sentence, first locate the verb 13/62

13. Nouns and pronouns may be used as the subject of a verb , the object of a verb , an indirect object , the object of a preposition , or a subject complement 44/62

14. If a group of words is to be a sentence , it must contain at least a subject and a predicate (verb) and express a complete thought. 6/62

15. If two nouns follow a verb of action, the first is usually an indirect object and the second, a direct object 19/62

16. A subject complement follows a linking verb and is equal to the subject or describes the subject 40/62

17. Write complete sentences to fit the following simple sentence patterns:
subject (noun) - compound verb (verb and verb) 13/62

18. subject (noun and noun) - verb - object . . . 14/62

19. subject - linking verb - complement (adjective and adjective) 25/62

20. subject - verb (adverb and adverb) 32/62

Median: 20/62

Range: . . . . 3/62 - 44/62

Appendix C

All of the lessons completed by each subject provided detailed records which were reduced to description in terms of the number of frames providing at least one kind of observation relative to the number of frames contained in that lesson. For any one of the five kinds of latency, there may be an additional missing observation or so. Averages are given for each kind of latency in terms of the actual number of observations contributing to that average. The range, lowest and highest latency value, is given for each kind of latency. All latency values are given in seconds.

From the left to the right in the first row or line of the table is given first the identifying number of the subject, the lesson and unit, the number of frames in the lesson portion providing some data / the total number of frames in the lesson portion of that lesson (i.e., review test frames are not included), the average value in seconds for each of the five kinds of latency, Read, Answer, Read-plus-Answer, Score-&-Turn, and Total Frame time. Below these averages, in the second line, are given the ranges.

Each error is considered in terms of the latency value associated with it. The average for each kind of latency is taken into account. If the value of a particular kind of latency associated with an error is greater than that kind of average, that error is classified Error Above Average (EAA). If it is below average it is classified EBA. The totals for both kinds of error are given as EAA/EBA. Adding numerator to denominator should give the total number of errors, although sometimes there may be a missing associated latency value.

Some Ss will have records for more than one lesson, some only one. The first table gives data reduced to the statistical categories discussed above as far as a first presentation of a lesson is concerned. The next table is similarly arranged and deals with a second presentation of a lesson. Finally, the third table deals only with the first presentation of the review test portion of some of the lessons of Unit II. All three tables provide the same kind of information and have the same format.

Since all interpretable data has been included, an occasional record obtained early in the study may give only Read-plus-Answer values or simply Total Frame.

Appendix C

Table 16

Range and Average Response Latencies and Errors Above Average Compared with Errors Below Average on First Presentation

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
1	1-II	65/67	Average:	12.75	14.17	26.92	5.92	32.85
			Range:	3-40	0-42	11-80	1-24	14-88
			EAA/EBA:	6/0	5/1	6/0	6/0	6/0
1	3-II	56/66	Average:	16.73	13.55	30.29	4.96	35.25
			Range:	2-55	3-73	9-100	2-23	12-117
			EAA/EBA:	4/3	4/3	5/2	7/0	5/2
2	3-II	35/66	Average:	11.86	7.31	19.17	4.94	24.11
			Range:	1-61	2-24	5-64	2-36	8-72
			EAA/EBA:	3/0	0/3	3/0	3/0	3/0
2	4-II	72/72	Average:	13.83	9.56	23.39	3.74	27.12
			Range:	3-45	2-35	5-67	2-11	7-71
			EAA/EBA:	6/3	3/6	4/5	9/0	5/4
3	1-II	67/67	Average:	----	----	16.71	----	----
			Range:	----	----	5-95	----	----
			EAA/EBA:	----	----	2/0	----	----
3	2-II	50/52	Average:	14.56	6.84	21.40	3.36	24.76
			Range:	3-59	0-31	4-67	1-8	7-70
			EAA/EBA:	0/0	0/0	0/0	0/0	0/0
3	3-II	63/66	Average:	17.05	11.25	28.30	4.44	32.75
			Range:	2-63	0-103	9-166	2-27	12-193
			EAA/EBA:	2/1	0/3	0/3	3/0	0/3
3	4-II	70/72	Average:	13.53	9.24	22.77	2.94	25.71
			Range:	3-34	0-35	7-46	1-8	9-49
			EAA/EBA:	2/0	0/2	0/2	2/0	1/1
3	5-II	52/52	Average:	9.67	12.06	21.73	3.23	24.96
			Range:	2-32	0-45	7-58	2-12	9-61
			EAA/EBA:	0/1	0/1	0/1	1/0	0/1

(Table continued on next page)

Table 16 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
4	1-IV	71/100	Average:	13.37	11.03	24.26	5.42	29.25
			Range:	4-33	0-32	9-57	3-19	13-62
			EAA/EBA:	1/0	0/1	0/1	1/0	1/0
4	5-VII	79/80	Average:	----	----	21.14	7.56	28.67
			Range:	----	----	9-75	4-17	14-85
			EAA/EBA:	----	----	2/1	3/0	3/0
4	6-VII[a]	23/70	Average:	10.48	9.65	20.13	9.65	26.96
			Range:	5-20	3-27	9-36	3-46	13-78
			EAA/EBA:	0/1	1/0	1/0	1/0	1/0
4	7-VII	80/80	Average:	12.01	11.73	23.74	6.64	30.38
			Range:	3-42	2-47	6-77	3-44	9-105
			EAA/EBA:	5/5	4/6	6/4	10/0	8/2
5	1-IV[a]	70/100	Average:	16.37	11.78	30.27	5.33	33.11
			Range:	4-48	2-51	11-79	2-23	14-65
			EAA/EBA:	6/4	4/6	5/5	9/0[b]	8/1[b]
5	2-IV	50/50	Average:	14.88	7.16	21.96	5.38	27.37
			Range:	2-65	1-22	5-66	2-21	7-81
			EAA/EBA:	6/2	4/4	6/2	6/2	7/1
5	5-VII	80/80	Average:	----	----	16.40	----	----
			Range:	----	----	4-47	----	----
			EAA/EBA:	----	----	17/15	----	----
5	6-VII	25/70	Average:	----	----	----	----	18.00
			Range:	----	----	----	----	8-46
			EAA/EBA:	----	----	----	----	5/0
5	7-VII	79/80	Average:	11.68	17.03	28.70	8.52	37.23
			Range:	1-41	2-95	6-121	2-53	10-165
			EAA/EBA:	6/6	7/5	6/6	9/3	7/5

---

[a] These data were the first successfully recorded although S was exposed to lesson two weeks earlier. (A successful repeat two days later yielded data considered Second Presentation.)

[b] Latency value associated with error was lost due to recording failure.

(Table continued on next page)

Table 16 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
6	1-II	60/67	Average:	9.65	8.07	17.72	4.30	22.02
			Range:	2-24	0-37	5-48	2-17	8-53
			EAA/EBA:	5/4	3/6	8/1	9/0	8/1
6	2-II	49/52	Average:	12.78	12.90	25.67	4.33	30.00
			Range:	1-33	0-54	10-70	2-15	12-74
			EAA/EBA:	1/2	3/0	2/1	3/0	2/1
6	3-II	62/66	Average:	13.56	8.71	22.27	4.24	26.52
			Range:	3-44	0-44	6-75	2-18	8-80
			EAA/EBA:	6/0	3/3	5/1	5/1	6/0
6	4-II	72/72	Average:	17.67	7.69	25.36	3.64	29.00
			Range:	5-61	0-35	8-75	1-12	10-79
			EAA/EBA:	3/2	2/3	2/3	5/0	5/0
6	5-II	52/52	Average:	10.15	10.15	20.31	3.98	24.29
			Range:	2-44	0-41	8-50	1-18	9-57
			EAA/EBA:	0/5	2/3	5/0	1/4	3/2
6	1-VII	80/80	Average:	10.59	4.06	14.65	2.85	17.50
			Range:	3-29	0-16	4-40	1-9	5-43
			EAA/EBA:	3/9	6/6	5/7	11/1	6/6
6	2-VII	78/79	Average:	11.28	4.08	15.36	3.38	18.74
			Range:	3-28	0-23	5-30	1-17	8-45
			EAA/EBA:	7/3	2/8	5/5	9/1	6/4
6	3-VII	45/45	Average:	12.42	4.33	15.39	3.78	18.86
			Range:	3-34	0-16	5-34	1-11	7-39
			EAA/EBA:	7/6	8/5	8/5	12/1	8/5
7	3-II	46/66	Average:	12.09	9.60	21.68	5.47	27.15
			Range:	0-37	2-35	8-66	2-24	11.78
			EAA/EBA:	4/1	1/4	4/1	5/0	4/1
7	4-II	72/72	Average:	11.33	----	----	3.96	----
Special record;			Range:	3-32	----	----	2-14	----
S did not write answers.								
7	5-II	52/52	Average:	14.85	17.38	32.23	4.98	37.21
			Range:	0-41	2-53	3-83	2-18	5-101
			EAA/EBA:	4/2	1/5	1/5	6/0	3/3

(Table continued on next page)

Table 16 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
7	2-VII	79/79	Average:	13.55	11.39	24.94	6.42	31.35
			Range:	5-29	4-34	11-52	2-29	17-78
			EAA/EBA:	2/3	3/2	2/3	5/0	4/1
7	3-VII	45/45	Average:	14.00	11.94	25.93	5.09	31.02
			Range:	3-37	3-28	7-61	2-22	9-64
			EAA/EBA:	7/6	8/5	8/5	12/1	8/5
8	5-VII	80/80	Average:	16.16	8.83	24.99	4.54	29.52
			Range:	4-60	2-44	7-79	2-17	10-91
			EAA/EBA:	12/16	10/18	12/16	23/5	15/13
8	6-VII	38/70	Average:	22.11	8.26	30.16	5.13	34.95
			Range:	4-100	2-23	8-122	2-18	12-129
			EAA/EBA:	5/5	6/4	6/4	6/4	8/2
8	7-VII	60/80	Average:	13.12	9.18	22.30	3.23	25.53
			Range:	1-41	0-41	5-48	1-10	7-57
			EAA/EBA:	6/3	3/6	4/5	7/2	4/5
9	1-II	64/67	Average:	15.16	8.11	23.27	5.67	28.94
			Range:	1-75	0-35	10-76	3-18	15-82
			EAA/EBA:	3/3	2/4	3/3	6/0	4/2
9	2-II	24/52	Average:	17.50	8.08	25.58	3.38	28.96
			Range:	5-49	0-22	10-52	2-10	14-54
			EAA/EBA:	1/0	0/1	1/0	1/0	1/0
9	3-II	66/66	Average:	14.28	7.63	21.95	3.22	25.20
			Range:	4-39	0-40	9-59	2-10	11-63
			EAA/EBA:	2/1	2/1	3/0	3/0	3/0
10	1-II	20/67	Average:	14.10	11.00	25.10	6.65	31.75
			Range:	4-31	3-34	8-69	3-17	12-80
			EAA/EBA:	2/1	0/3	2/1	2/1	2/1
10	2-II	52/52	Average:	15.23	11.69	26.92	6.24	31.82
			Range:	3-54	0-103	6-120	3-34	11-130
			EAA/EBA:	1/3	0/4	1/3	3/1	1/3
11	1-IV	100/100	Average:	17.90	13.33	30.40	3.94	35.17
			Range:	4-75	1-64	6-98	1-19	8-116
			EAA/EBA:	11/15	12/14	10/16	24/2	12/14

(Table continued on next page)

Table 16 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
11	2-IV	48/50	Average:	13.76	9.98	22.73	3.63	26.37
			Range:	4-35	1-64	6-98	1-19	8-116
			EAA/EBA:	2/0	2/0	2/0	2/0	2/0
11	5-VII	78/80	Average:	----	----	22.23	5.50	27.73
			Range:	----	----	6-75	2-44	8-107
			EAA/EBA:	----	----	9/5	11/3	9/5
11	6-VII	69/70	Average:	----	----	22.36	4.42	26.93
			Range:	----	----	6-101	1-32	8-124
			EAA/EBA:	----	----	4/7	11/0	8/3
11	7-VII	80/80	Average:	11.99	9.68	21.66	3.83	25.48
			Range:	2-45	0-42	5-72	1-23	7-95
			EAA/EBA:	12/1	8/5	10/3	13/0	11/2
12	1-II	67/67	Average:	14.80	16.31	30.88	7.28	38.15
			Range:	2-69	0-78	7-120	3-40	11-132
			EAA/EBA:	6/4[b]	8/3	9/1[b]	10/1	9/1[b]
12	1-VII	42/80	Average:	26.93	14.12	41.20	11.48	52.88
			Range:	4-149	0-44	11-149	3-75	14-186
			EAA/EBA:	7/5	3/9	5/7	8/4	5/7
13	1-II	46/67	Average:	10.78	18.10	28.87	7.02	35.89
			Range:	3-31	3-74	6-105	3-27	11-115
			EAA/EBA:	6/4	7/3	7/3	9/1	7/3
14	2-II	52/52	Average:	23.85	10.58	34.42	5.15	39.58
			Range:	0-98	0-38	10-98	2-25	13-109
			EAA/EBA:	4/2	2/4	4/2	5/1	5/1
15	2-II	52/52	Average:	----	----	18.80	3.88	22.73
			Range:	----	----	4-57	2-19	7-66
			EAA/EBA:	----	----	2/0	2/0	2/0
15	3-II	60/66	Average:	12.70	7.27	19.97	4.75	24.72
			Range:	0-36	0-25	7-60	2-20	10-65
			EAA/EBA:	4/7	6/5	8/3	9/2	8/3
16	5-II	52/52	Average:	9.04	14.25	23.29	4.60	27.88
			Range:	1-31	0-41	6-58	3-12	11-63
			EAA/EBA:	2/2	2/2	3/1	4/0	3/1

---

[b] Latency value associated with error was lost due to recording failure.

Table 17

Range and Average Response Latencies and Errors Above Average

Compared with Errors Below Average on Second Presentation

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
3	1-II	67/67	Average:	----	----	15.79	----	----
			Range:	----	----	4-43	----	----
			EAA/EBA:	----	----	1/0	----	----
3	5-II	52/52	Average:	8.19	14.25	23.29	4.60	27.88
			Range:	2-21	0-34	5-48	2-15	7-51
			EAA/EBA:	0/0	0/0	0/0	0/0	0/0
4	5-VII	36/80	Average:	10.44	9.26	19.71	6.15	25.85
			Range:	5-33	2-35	11-47	2-13	17-50
			EAA/EBA:	0/0	0/0	0/0	0/0	0/0
5	1-IV	70/100	Average:	8.58	11.26	19.67	5.05	23.44
			Range:	3-22	4-31	7-51	2-26	11-62
			EAA/EBA:	10/2	7/5	8/4	12/0	10/2
5	2-IV	50/50	Average:	14.82	9.10	23.84	4.68	28.53
			Range:	2-43	1-40	6-53	2/43	8-91
			EAA/EBA:	2/3	5/0	5/0	5/0	5/0
5	5-VII	79/80	Average:	----	----	15.90	----	----
			Range:	----	----	5-47	----	----
			EAA/EBA:	----	----	9/12	----	----
5	6-VII	70/70	Average:	----	----	20.17	4.73	24.90
			Range:	----	----	7-70	2-15	10-78
			EAA/EBA:	----	----	9/5	14/0	12/2
5	7-VII	45/80	Average:	8.58	11.27	19.84	4.02	23.87
			Range:	1-20	2-29	7-41	2-12	9-47
			EAA/EBA:	2/0	0/2	0/2	1/1	1/1
6	4-II	72/72	Average:	9.93	8.85	18.78	3.22	22.00
			Range:	1-41	2-30	7-51	1-18	8-53
			EAA/EBA:	3/1	1/3	3/1	4/0	3/1
6	1-VII	80/80	Average:	9.66	3.46	13.13	2.71	15.84
			Range:	1-38	0-12	3-42	1-9	5-44
			EAA/EBA:	4/3	1/6	5/2	7/0	5/2

(Table continued on next page)

Table 17 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
7	2-VII	31/79	Average:	8.65	8.65	16.32	3.65	21.26
			Range:	3-33	3-26	9-44	2-11	11-49
			EAA/EBA:	0/3	0/3	0/3	3/0	2/1
8	5-VII	80/80	Average:	14.55	7.93	22.48	4.00	26.50
			Range:	1-98	0-44	7-117	2-19	9-133
			EAA/EBA:	13/9	7/15	13/9	18/4	13/9
8	6-VII	38/70	Average:	17.11	8.79	25.73	6.08	31.73
			Range:	6-39	2-38	10-76	2-21	13-90
			EAA/EBA:	8/2	7/3	9/1	8/2	10/0
10	2-II	52/52	Average:	11.54	10.65	22.19	4.88	26.12
			Range:	2-40	1-68	6-94	2-22	10-105
			EAA/EBA:	2/2	2/2	3/1	4/0	3/1
11	1-IV	70/100	Average:	16.99	10.67	27.66	4.06	31.71
			Range:	5-72	2-51	9-86	1-38	12-88
			EAA/EBA:	5/3	3/5	5/3	7/1	5/3
11	5-VII	78/80	Average:	9.46	7.44	17.01	2.99	20.00
			Range:	3-37	0-51	6-59	1-28	7-65
			EAA/EBA:	5/6	4/7	5/6	11/0	5/6
11	6-VII	69/70	Average:	11.17	9.26	20.43	3.67	24.10
			Range:	4-41	2-41	9-78	1-32	11-95
			EAA/EBA:	4/3	6/1	6/1	7/0	7/0
11	7-VII	80/80	Average:	9.40	9.13	18.53	3.31	21.84
			Range:	2-50	1-27	4-52	1-15	6-55
			EAA/EBA:	6/4	7/3	7/3	10/0	10/0
15	3-II	60/66	Average:	9.38	5.95	15.33	4.18	19.52
			Range:	1-25	0-28	4-39	2-19	7-43
			EAA/EBA:	2/2	3/1	3/1	4/0	4/0

Table 18

Review Test Range and Average Response Latencies and E-Scored Errors Above Average Compared with Errors Below Average on First Presentation

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
1	1-II	20/20	Average:	17.75	19.85	37.60	4.55	42.15
			Range:	3-122	0-49	12-122	2-16	16-125
			EAA/EBA:	5/3	6/2	4/4	3/5	4/4
1	3-II	20/20	Average:	18.20	26.00	44.20	2.60	46.80
			Range:	4-51	2-81	12-120	1-4	14-122
			EAA/EBA:	4/3	2/5	1/6	5/2	1/6
3	1-II	19/20	Average:	----	----	30.21	----	----
			Range:	----	----	7-95	----	----
			EAA/EBA:	----	----	7/2	----	----
3	2-II	16/20	Average:	19.50	4.75	24.25	2.94	27.19
			Range:	6-45	0-27	8-45	1-5	9-49
			EAA/EBA:	3/2	1/4	3/2	4/1	4/1
3	5-II	20/20	Average:	17.10	22.55	39.65	2.35	42.00
			Range:	3-37	0-88	8-114	1-6	10-116
			EAA/EBA:	7/2	4/5	6/3	2/7	4/5
6	3-II	20/20	Average:	16.75	16.30	33.05	2.55	35.60
			Range:	4-40	0-43	10-59	2-5	12-61
			EAA/EBA:	4/5	5/4	6/3	4/5	6/3
6	5-II	19/20	Average:	17.89	17.89	35.79	2.53	38.32
			Range:	3-51	0-81	8-97	1-10	10-100
			EAA/EBA:	5/7	6/6	6/6	5/7	7/5
7	3-II	19/20	Average:	11.85	22.05	33.90	4.00	37.90
			Range:	0-46	6-45	9-72	3-7	16-76
			EAA/EBA:	3/1	3/1	4/0	4/0	4/0
7	5-II	17/20	Average:	13.71	47.24	60.94	2.47	63.41
			Range:	0-57	8-242	11-299	2-4	14-301
			EAA/EBA:	2/4	4/2	3/3	2/4	3/3

(Table continued on next page)

Table 18 (Continued)

S#	Lesson	Frames	Statistical Categories	Read	Answer	R-plus-A	S-&-T	Total Frame
9	3-II	20/20	Average:	15.55	8.40	23.95	3.10	27.05
			Range:	5-45	0-15	9-45	1-10	12-47
			EAA/EBA:	3/3	4/2	4/2	3/3	4/2
10	1-II	19/20	Average:	16.67	18.22	34.89	4.56	39.44
			Range:	5-53	2-46	15-57	3-8	18-61
			EAA/EBA:	2/3	3/2	4/1	2/3	4/1
10	2-II	20/20	Average:	19.55	22.25	42.30	4.05	46.35
			Range:	4-112	0-98	10-134	3-10	13-137
			EAA/EBA:	3/3	3/3	3/3	0/6	3/3
13	1-II	19/20	Average:	14.21	46.05	60.26	4.10	64.47
			Range:	2-47	8-100	10-133	1-8	14-141
			EAA/EBA:	4/4	3/5	3/5	5/3	3/5
15	2-II	19/20	Average:	----	----	29.95	3.00	32.95
			Range:	----	----	5-65	2-8	7-68
			EAA/EBA:	----	----	5/7	9/3	6/6
15	3-II	20/20	Average:	18.35	8.20	26.55	3.30	29.85
			Range:	3-49	0-29	7-52	2-6	10-55
			EAA/EBA:	5/5	6/4	8/2	3/7	8/2
16	5-II	17/20	Average:	11.35	25.82	37.18	3.00	40.18
			Range:	1-26	0-85	5-98	2-5	8-101
			EAA/EBA:	5/2	3/4	3/4	6/1	3/4

VITA

Lloyd Oliver Brooks

Personal History: Born March 7, 1927, Lincoln, Missouri

Father, Samuel Brooks (deceased) Mother, Eva Williams Brooks (deceased)

Wife, Polly H. (MacLeod) Brooks - Married 1955

Education:

1941-1945 - Graduated from West Plains High (Missouri)
(1945-1946 - U. S. Navy)

1949-1952 - Undergraduate work: Long Beach City College (California) University of Florida Extension, Florida Southern (Lakeland, Florida) and Long Beach State College (California)

1952-1953 - University of California at Los Angeles Received B.A. Major: Psychology

1954-1958 - Study toward Ph.D. degree in Experimental Psychology, U.C.L.A. Graduate Division

1958-1961 - This study continued at the University of Houston.

Experience:

1953-1957 - Research Assistant - U.C.L.A.

1958 - Laboratory Assistant-Research Assistant U. of H.

1958-1960 - Teaching Fellow - U. of H.

1960-1961 - Research Assistant - U. of H.

Permanent Address:

Department of Psychology
Arizona State University
Tempe, Arizona